AD NUMBER: AD0896849

LIMITATION CHANGES

TO:

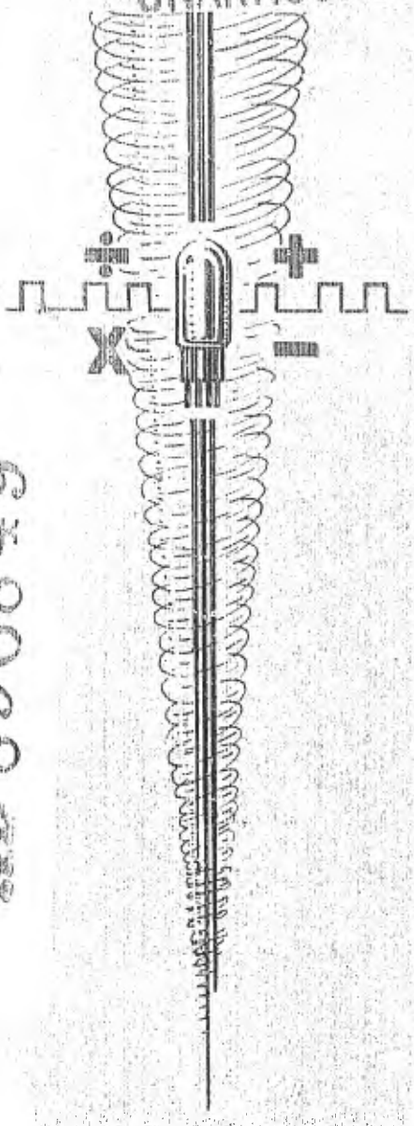
Approved for public release; distribution is unlimited.

FROM:

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; 09 NOV 1977 Other requests shall be referred to Office of Naval Research, Arlington, VA 22203.

AUTHORITY

ONR ltr dtd 09 NOV 1977



PROJECT WHIRLWIND

Contract N5ori60

SUMMARY REPORT NO. 2

VOLUME 8

MATHEMATICS

SERVOMECHANISMS LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SPECIAL DEVICES CENTER

UNANNOUNCED

(1)

M-142

PROJECT WHIRLWIND

(9) Summary Report No. 2.

(11) November 1947

(15) N5 ori-60

(12) 65p.

(6) Project Whirlwind.
Volume 8.

MATHEMATICS.

Volume 8 of 22 Volumes

Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

220 076

CONTENTS

INTRODUCTION

The original mathematical work on the project was started in the early part of 1946, and constituted an investigation of methods of interpolation and integration to be used in the solution of the airplane stability control problem. There are no published reports on this study but the results were used in the programming of the ASCA problem which is reported in Note C-15 in Vol. 14.

The project was later ~~joined by Dr. Loud who began~~ engaged in an investigation of the stability of the numerical solutions proposed for the airplane problem. This study included investigations of simple linear and non-linear systems, and considered the use of mathematical filters and energy conservation methods to avoid instability. Some of these studies were reported in M-32 and in R-114. In this work he was assisted by Margaret Florencourt, Kenneth B. Tuttle, and Edgar Reich.

Starting in March of 1947, a series of meetings were held between the mathematics people of Project Whirlwind and the Mathematics Department of M.I.T. M-61 by Dr. Loud presents some ideas for a course in numerical methods discussed during these meetings. Professor Norbert Wiener of M.I.T. was present at these meetings and presented some important ideas, two of which are discussed by Dr. Loud in M-71 and M-78. Later, a study was begun on the application of Wiener's statistical filter theory to the problems of numerical solutions. A part of this work is reported by Dr. Loud in M-94 and M-95.

Dr. Loud left the project for another university during the summer of 1947. The mathematics work on the project was then put under the guidance of Professor Franklin of M.I.T.'s Mathematics Department. As part of the work since that time, an investigation has been made of probability methods useful in the correlation of radar and other data useful in such problems as the airport control problem. Part of this work is reported by Professor Franklin in M-124. M-100 is a trip report on a computing machine symposium attended by Edgar Reich.

The present mathematical work of Project Whirlwind as well as plans for future work are discussed by Professor Franklin in M-160.

REFERENCE INDEX

M Series Memorandums

REF.	VOL.	REF.	VOL.	REF.	VOL.
M-32	8	M-95	8	M-133	18
M-46	9	M-96	9	M-134	7
M-56	9	M-99	15	M-135	7
M-58	15	M-100	8	M-136	7
M-61	8	M-101	11	M-137	7
M-62	4	M-103	16	M-138	15
M-63	4	M-105	19	M-140	4
M-64	4	M-106	11	M-141	7
M-65	14	M-107	19	M-142	8
M-66	4	M-109	16	M-143	9
M-68	15	M-110	15	M-144	10
M-69	4	M-111	7	M-145	11
M-71	8	M-112	9	M-146	12
M-72	16	M-113	7	M-147	13
M-74	14	M-114	19	M-148	14
M-76	4	M-116	16	M-149	15
M-77	15	M-117	7	M-150	16
M-78	8	M-118	16	M-151	17
M-80	16	M-119	16	M-152	18
M-81	16	M-121	9	M-153	19
M-82	16	M-123	7	M-154	20
M-83	16	M-124	8	M-155	21
M-85	14	M-127	7	M-156	22
M-89	11	M-128	16	M-157	11
M-91	15	M-129	7	M-158	7
M-92	15	M-130	9	M-159	9
M-94	8	M-131	16	M-160	8
		M-132	16	M-161	7

REFERENCE INDEX

E Series Memorandums

C Series Memorandum

REF.	VOL.	REF.	VOL.
E-7	14	E-52	19
E-24	7	E-53	13
E-31	10	E-54	19
E-32	10	E-55	19
E-33	19	E-56	15
E-37	15	E-57	15
E-38	19	E-58	19
E-39	15	E-59	19
E-41	15	E-60	19
E-42	15	E-61	16
E-44	19	E-63	19
E-45	19	E-64	15
E-47	15	E-68	13
E-48	19	E-69	16
E-49	19	E-71	19
E-50	16	E-73	16
C-15	14		

REFERENCE INDEX

R Series Memorandums

REF.	VOL.	REF.	VOL.
R-36	14	R-115	4
R-49	14	R-116	4
R-63	14	R-117	16
R-64	3	R-118	16
R-89	19	R-120	10
R-90	4	R-121	19
R-94	14	R-122	18
R-96	14	R-123	17
R-100	14	R-124	11
R-103	14	R-125	14
R-104	16	R-126	19
R-106	15	R-127	5
R-108	15	R-127	6
R-109	19	R-128	10
R-110	9	R-129	12
R-111	15	R-130	9
R-113	15	R-131	10
R-114	8	R-133	10

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: MATHEMATICAL WORK OF PROJECT WHIRLWIND

To: Jay W. Forrester

From: Philip Franklin

Date: November 12, 1947

I. The Mathematics Program

The projected Whirlwind high-speed electronic digital computer, from the mathematical point of view, has several objectives. One aim is to solve the equations of motion for aircraft and so serve as the computing element of an aircraft analyzer. A more general objective is the solution of other problems of engineering interest, particularly in the field of dynamics and ordinary differential equations.

The mathematical studies in the past have related to the specific equations to be solved, some tests of existing numerical methods, and attempts to anticipate possible difficulties with a view to prescribing how to meet them. At present a study of known methods of solving simultaneous linear equations and differential equations is under way, with a view to the coding of such methods for the computer. These studies will be extended to other types of problems, such as those listed in Section III.

II. Studies Already Completed

A survey of numerical methods made by Dr. Loud and his associates, summarized in Memorandum M-61 (Vol. 8), revealed as subjects of major importance the solution of linear simultaneous equations, and the solution of ordinary differential equations.

For the solution of linear simultaneous equations, elimination methods, iteration methods, relaxation methods, matrix methods, and the method of steepest descent have all been considered and expositions of these methods studied. For some of these, codes are in process of preparation.

The errors inherent in solving systems of equations of very high order, when all of the coefficients have random errors, has been pointed out.

A number of solutions have been carried out for simple differential equations to provide case histories of accumulated round-off and truncation errors. These indicate the objections to numerical integration over a large

number of cycles without some check, such as the use of the energy integral for conservative systems. The possibility of filtering out frequencies corresponding to computational instability terms has also been considered.

A method of using Fourier Transforms to estimate the best value of a function or its derivative (the message) when its values over a long range are known only to within certain random errors (noise) was presented to the group. Later studies indicated this as having restricted application to computation because of the rare occurrence of data with truly random errors.

Last year several conferences were held at M. I. T. on digital computation. Some were sessions of the Electrical Engineering Department seminar, but were well attended by representatives of the Mathematics Department and Project Whirlwind. Another series of smaller conferences were attended by some electrical engineers, including Professors Hazen and Y. W. Lee, as well as several members of the Center of Analysis. The mathematicians present included Professors Phillips, Martin, Wiener, Franklin, Wallman, and Thomas.

In addition to these men, several other M. I. T. staff members have shown an interest in electronic computers, and a willingness to give advice when problems in their special fields are considered. Thus, if shock waves are under consideration, advice can be obtained from Professors Tsien or Lin or Dr. Kopal, who are all familiar with this field. For elasticity problems and their theory Professor Reissner, for statistical and meteorological applications Professor Wadsworth, and for general numerical computation Professors Hitchcock, Crout, and Hildebrand are available.

The many members of the M. I. T. Mathematics Department who are interested in numerical computation contribute much information about recent developments through informal discussions.

In addition, the staff members of the project have attended meetings at Princeton, Philadelphia, New Haven, and Cambridge at which computing methods were discussed. And the group has followed recent developments in the theory of codes and numerical computation at Princeton, Harvard, the University of Pennsylvania, and elsewhere.

III. Plan for Future Investigations

Certain problems of interest to applied scientists and engineers can be most conveniently solved by numerical methods. Classified mathematically, the principal types are as follows:

1. Solution of ordinary differential equations with given initial conditions.

2. Solution of parabolic and hyperbolic partial differential equations, obtaining the characteristics curves as in 1.

3. Solution of systems of linear simultaneous equations.

4. Solution of ordinary differential equations with boundary conditions at more than one point.

5. Solution of elliptic partial differential equations for various types of conditions on the boundary of a region.

6. Solution of non-linear simultaneous equations.

7. Least square solutions of overdetermined systems of the types of 3 and 6.

8. Solution of integral equations.

9. Tabulation of functions.

While the relative advantages of different methods may change in the transition from hand or desk calculations to high-speed machines, it is improbable that any radically new mathematical principles will come into play. Much past and recent experience bears this out.

Thus, the methods used by Aiken in computing tables, or by the Watson Laboratory in checking the moon's motion, bear strong resemblance to those long known to astronomers; and except for the use of conditional orders the methods used for elementary functions, such as those coded for the A.R.C. by Booth and Britten, would have been used by Babbage in the nineteenth century if he had completed his analytical engine.

Again, the report on linear equations by Bargmann, Montgomery, and von Neumann recommends as the two best methods an iteration rule, stated by Hotelling, which amounts to an application of Newton's method of approximation, or an elimination method which is that used by most computers in such forms as those of Doolittle or Crout.

And except for minor details of technique, the differential analyzer uses mathematical methods differing little from those used by Kelvin in evaluating special integrals with his globe disk and cylinder integrator.

This suggests that the first attack with the computer on the problems listed above should be by processes close to the traditional ones. Of course, two special points to be covered in taking over existing methods are the coding of complete instructions to remove any human judgment, and keeping the high speed from letting us carry the computations beyond their range of validity.

The present thinking on procedures for the various types of problems is as follows:

1. Ordinary Differential Equations, One Point Conditions

For systems of ordinary differential equations, some method using second differences or their equivalent such as the

Runge-Kutta method or the Moulton (Adams, Krilloff) method. This last may be used in the form set up by Ford, which uses linear combinations of the numbers instead of differences. After several steps have been taken, the results can be checked and improved by a Simpson's rule calculation which would have a smaller round-off error than the result obtained step by step.

2. Characteristics of Partial Differential Equations

Most of the procedure is as in 1, the new feature being the cut-off due to an initial boundary or the meeting of some conditional bounding condition like that for shock waves.

3. Linear Algebraic Equations

For systems in 20 or fewer variables, an elimination method can be used and coded with about 80 orders. The matrix iteration procedure $F_{k+1} = F(2-AF_k)$ to obtain A^{-1}, the limit of F_k, may be useful for more variables or where a system must be solved for several right members. A steepest-descent method on the sum of the squares of the residuals may be useful for a large number of variables.

In many applications where numerous variables are met, the system is loosely coupled in the sense that effects in one place affect values at a distance only slightly. Here much of the difficulty of treating a large matrix with all values equally in error disappears.

Systems with determinants near zero will have to be recognized as such at some stage of the solution process, and the recognition used as a halt signal.

4. Ordinary Differential Equations, Boundary Conditions

This situation will be met by reduction to an integral equation, by an expansion in known functions whose co-efficients are found by solving linear equations (Ritz, Galerkin, Fourier), or from solutions found as in 1. For conditions at two points we use a family of solutions starting from one point, and interpolate to meet the conditions at the second point. For linear differential equations with conditions at several points, we may use a linear combination of independent solutions found numerically, evaluating the constants by solving a system of linear algebraic equations.

5. Elliptic Partial Differential Equations

Linear partial differential equations of elliptic type are reduced to difference equations on nets. In a two-dimensional problem, if our machine eventually used say 64 points on a square

it might be advantageous to get a first approximation for a 4 x 4 grid by solving simultaneous equations, then interpolating on this for a first approximation on an 8 x 8 grid, which could be improved by some successive approximation scheme or relaxation method. Relaxation methods by machine have the difficulty of requiring a large number of comparison orders.

WWI will probably be able to work two-dimensional grids of fairly fine dimensions, but possibly only crude three-dimensional grids, like those now used by hand in two-dimensional problems.

6. Non-Linear Algebraic Equations

Iteration procedures are usually successful. The code would have to make possible the insertion of an approximation at different points in a loop, only continuing if the successive values clustered after a few tries.

7. Least Squares

Gauss procedure reduces this to 6 and 3.

8. Integral Equations

Approximation of the kernel by a degenerate one reduces linear integral equations to 3, or we can express the solution as a linear combination of known or determined functions as described under 4 above.

9. Functions

Our machine will probably tabulate functions only for its own use. Functions will be inserted either through discrete values and an interpolation program, or by using polynomials that approximate the functions in restricted ranges and calculating the polynomials.

In some cases the approximating polynomials may be found by approximating some higher derivative and integrating.

In addition to general methods, several aspects of the specific application of the computer to the analysis and control problem will be considered. Thus, the detailed equations and constants for different types of service and air-borne craft will be investigated.

Some preliminary studies of the mathematics of correlating statistically, at information centers, data from radar stations will be continued.

IV. The Practical Estimate of Error

While existence theorems are nice to have when they can be obtained,

much numerical work proceeds successfully without them. For example astronomers guessed right on semi-convergent asymptotic series before they were understood by the purists. Series with coefficients obtained by numerical methods are used to predict the moon's position years in advance, but the stability of the system of sun, earth, and moon under the known initial conditions has never been established with complete mathematical rigor. Many of the iteration processes used to find roots of equations can either converge or diverge rapidly. By comparing, say, a fourth and a fifth approximation, and sensing the size of the discrepancy, our machine will recognize the process as finished, promising for further approximations, or useless for computation because of divergence or excessively slow convergence.

A large percentage of recent numerical computation for partial differential equations has no other justification than that the results seem to converge. Undoubtedly many WWI and WWII solutions will use such evidence, for lack of anything better.

Because of the labor of making codes, high-speed machines will be chiefly used for whole fields of solutions. A certain amount of experiment on the machine may be needed at the start, but once verified for a few cases, the method will extend to the whole field.

Similar considerations apply to formulas for the size of the error. These are seldom used literally, because if rigorous they are apt to be much too pessimistic in practice. At best such formulas are merely useful to prove convergence; a comparison of successive values then estimates the error.

In many problems the stability of the method of solution can be predicted in advance, and a large number of solutions differing slightly are to be found. Here the smooth character of the results is a reasonable protection against excessive errors at any place.

In other cases, some form of a successive approximation method will be used. When the process succeeds, the clustering of the approximations serves both to check the results and to provide an estimate of the closeness of approximation.

When the successive approximation method fails, some order to test the closeness of approximations after some set number of steps will lead either to a halt signal or to an alternative procedure. Thus in a loop process the alternative may be another starting point in the loop. In a step-by-step process the alternative may be the use of smaller steps.

In many solutions of ordinary differential equations by the Moulton method, the use of a Simpson's rule integration over large intervals may prevent the piling up of round-off errors.

V. The Influence of Mathematics on the Design of WWI

The specifications for WWI were set up with the aim of providing a prototype capable of performing all the operations desired from a computer, on a scale large enough to be applied to many typical situations, and sufficiently limited to be constructable in reasonable time. However, within this last limitation, it was found feasible to provide for greater speed and storage capacity than any existing digital computer possesses.

The types of desired mathematical operations are fairly well known: namely, the four fundamental arithmetic operations, and certain logical operations such as shifting, digit transfers, subprogramming, and conditional subprogramming. From these fundamental elements any finite combination of computational processes can be built up.

The mathematician's requirement is that these operations be performed either directly, or by combination of other operations. Thus division or square rooting can be either built in as special operations, or carried out by a subprogram. The present plans call for a built-in division unit, and the possibility of including a few special orders which would enable the operator to call for any particular subprogram, as that for square root.

Considerable study of binary arithmetic and the conversion problem showed the feasibility and advantage of having the computer work in binary arithmetic but translate data and outputs from or into the decimal notation.

A guiding principle of the WWI specifications has been maximum flexibility. This will simplify the extension to larger models, and will increase the effective power of the computer, for example by allowing an arbitrary distribution of the storage capacity between numbers and orders.

Beyond the fundamental questions just mentioned, and verification that the computer has sufficient speed and capacity for certain specific problems such as that of aircraft control, mathematical considerations can have little effect on the basic design.

With regard to overall capacity of later models, experience with WWI and theoretical considerations can only dictate ratios of components, rather than absolute size. For the size at any stage will always be limited by economic and engineering factors. Any machine is certain to have just a little less capacity than that required for some desired applications. For WWI, sufficient capacity for the aircraft problem seems to be the only fixed requirement; that is, it should have sufficient speed and capacity to be used as a computing link in the aircraft analyzer.

Philip Franklin

Philip Franklin

PF:hcs

MEMORANDUM NO. M-32

TO: J. W. Forrester

FROM: Warren S. Loud

SUBJECT: Numerical Solution Method

DATE: November 18, 1946

Suppose $\frac{dy}{dt} = f(y)$ and we wish to set up a numerical solution with a step-by-step process. Let the interval between steps be h and assume the values of y at 0, h, 2h, 3h, ---nh, are known. The value at (n+1) h will be found as follows:

1. Extrapolate the solution to (n+1)h by some means.

2. Integrate by some means: $\int_{nh}^{(n+1)h} f(y)\, dt$

3. Set $y_{n+1} = y_n + \int_{nh}^{(n+1)h} f(y)dt$

For small enough values of h, they will follow the true solution closely.

Notes:

1. The extrapolation procedure might be a linear process, i.e., extension of the solution from (nh, y_n) by a straight line. This may be the line through $\left[(n-1)h, y_{n-1}\right]$, $\left[nh, y_n\right]$ or it may be the line through (nh, y_n) with slope equal to $f(y_n)$.

Again we may use parabolic extrapolation using y_n, y_{n-1} and y_{n-2} or using y_n, y_{n-1} plus the slope at y_n.

2. The integration is the area under $f\left[y(t)\right]$ where $y(t)$ is the extrapolated curve. A linear or parabolic approximation is made.

The method was extended to a second order equation. Consider the equation

$$\ddot{y} = -y$$

If $\dot{y} = v$, $\dot{v} = -y$ and we have simultaneous first order equations. Let us use linear extrapolation and integration, using the point and the slope. The extrapolations are then:

$$y = y_n + v_n (t - nh)$$

$$v = v_n - y_n (t - nh)$$

The integrals from t = nh to (n+1) h of $\frac{dy}{dt}$ and $\frac{dv}{dt}$ are respectively

$$hv_n - \frac{h^2}{2} y_n \quad \text{and} \quad -hy_n - \frac{h^2}{2} v_n \quad \text{so that the new points are}$$

$$y_{n+1} = y_n (1 - \frac{h^2}{2}) + h v_n$$

$$v_{n+1} = v_n (1 - \frac{h^2}{2}) - h y_n$$

These equations can be solved to give

$$y_n = r^n \left[y_o \cos n \phi + v_o \sin n \phi \right]$$

$$v_n = r^n \left[-y_o \sin n \phi + v_o \cos n \phi \right]$$

$$\text{where} \quad r = \sqrt{1 + \frac{h^4}{4}} \quad \text{and} \quad \tan \phi = \frac{h}{1 - h^2/2}$$

If plotted, letting t = nh, and setting $y_o = 0$ $v_o = 1$ so that the true solution is y = sin t, we find all the points be on the curve

$$y = e^{\beta t} \sin wt \quad \text{where}$$

$$\beta = \frac{1}{2h} \log (1 + \frac{h^4}{4}) \approx \frac{h^3}{8} \qquad w = \frac{\phi}{h} \approx 1$$

Thus the approximate solution would be (if linear interpolation is used) a succession of chords of the curve

$$y = e^{\beta t} \sin wt.$$

Remarks:

1. If more complicated extrapolation and integration is used, the points are on a curve whose equation is of the form

$$y = e^{\beta_1 t} \sin w_1 t + e^{\beta_2 t} \sin w_2 t + \ldots$$

where for small h all the β's but one are large and negative.

2. It happens that for all methods tested the value of β is definitely positive for large values of h. This produces bad divergence. The worst values seem to come for $h \approx 3$, but $\beta > 0$ for $h > 2$ in most cases. In the method explicitly described, $\beta > 0$ for all values of h.

3. If damping was introduced into the equations, β still became positive for large h.

4. In a system with two natural frequencies, $\ddot{y} + 101\dot{y} + 100y = 0$ with $h = .1$ which is 20 points per period for the $w = 1$ component but only 2 points per period for the $w = 10$ component, the divergence was present and not suppressed due to the small number of steps per period in the high frequency part.

Warren S. Loud

Warren S. Loud

WSL;mla

6545
Report No. R-114

SERVOMECHANISMS LABORATORY
Massachusetts Institute of Technology
Cambridge, Massachusetts

Date of Report: February 24, 1947 Page 1 of 10 pages

Written by: Warren S. Loud

Subject: A Numerical Integration Method

Curves:
B-38146-G-1
B-38170-G
B-38171-G
B-38174-G-1

PART I - Description of the Method

In connection with the solution of the aircraft equations, it was thought desirable to study the various possibilities for step-by-step solution. The principal method under consideration might well be called the extrapolation and integration method. It is with the characteristics of this method that this report deals. The discussion is centered about its application to vibrating systems.

The following is the method for a first order equation. The type of equation handled is $\frac{dy}{dt} = f(y)$.

A step-by-step solution is set up as follows: Let h be a small interval in time, and assume the solution to be known at $t = 0,\ h,\ 2h,\ \cdots (n-1)h$. Let the values of y at these points be $y_0,\ y_1,\ y_2,\ \cdots,\ y_{n-1}$. The value of y_n is found by a two-part process. First, the function $f[y(t)]$ as a function of t is extrapolated from $t = (n-1)h$ to $t = nh$. Second, this function is integrated from $t = (n-1)h$ to $t = nh$; the value of the integral is added to y_{n-1} to obtain y_n. It is shown in the theory of differential equations that the true solution will be obtained within any desired approximation for some length of time provided that h is taken small enough.

The methods of extrapolation and integration may be chosen from a number of possibilities. The most usual methods are linear and parabolic. A linear extrapolation would be extension of $f[y(t)]$ by a straight line. This line could be the line through the points $\left[(n-2)h,\ f(y_{n-2})\right]$ and $\left[(n-1)h,\ f(y_{n-1})\right]$. It could also be the line through the point $\left[(n-1)h,\ f(y_{n-1})\right]$ having the slope $f'(y_{n-1}) \cdot f(y_{n-1})$. If the function $f(y)$ is simple(e.g. linear) this last is not difficult to compute and there is the advantage of simplicity in analysis. Similarly, a parabolic extrapolation would be extension of $f[y(t)]$ by a parabola, either that passing through $\left[(n-3)h,\ f(y_{n-3})\right]$, $\left[(n-2)h,\ f(y_{n-2})\right]$ and $\left[(n-1)h,\ f(y_{n-1})\right]$ or that passing through the second two of those and having the slope $f'(y_{n-1})\ f(y_{n-1})$ at the point for $t = (n-1)h$.

In certain cases the values y_0, y_1, y_2, etc., can be found explicitly. For example, let the equation be $\frac{dy}{dt} = Ay$, and let the extrapolation method be linear using the slope. The extrapolation will be:

$$\Delta y^* = (\Delta y_{n-1}) + \left[t - (n-1)h\right] \cdot \Delta (\Delta y_{n-1}).$$

Integrating from (n-1)h to nh and adding y_{n-1}, we obtain

$$y_n = y_{n-1} + h (\Delta y_{n-1}) + \frac{h^2}{2} \cdot \Delta (\Delta y_{n-1})$$

or,

$$y_n = y_{n-1} \left[1 + h\Delta + \frac{h^2 \Delta^2}{2}\right]$$

This difference equation has the solution

$$y_n = y_0 (1 + h\Delta + \frac{h^2\Delta^2}{2})^n$$

and if we replace n h by t, we obtain

$$y = y_0 \left\{\left[1 + h\Delta + \frac{h^2\Delta^2}{2}\right] h^{\frac{1}{h}}\right\}^{\Delta t}$$

so that the values y_0, $y_1 \cdots y_n \cdots$ lie on the curve

$$y = y_0 \left\{\left[1 + h\Delta + \frac{h^2\Delta^2}{2}\right] \frac{1}{h}\right\}^{\Delta t}$$

spaced a distance h apart in t. If $h \to 0$, since $\lim_{h \to 0} \left(1 + \Delta h + \frac{\Delta^2 h^2}{2}\right)^{\frac{1}{\Delta h}} = e$ we obtain $y = y_0 e^{\Delta t}$ which is the true solution.

PART IX - Application to Vibrating Systems

The principal study of the method described was its application to vibrating systems. The investigation was limited for the most part to the equation

$$\frac{d^2y}{dt^2} + y = 0$$

which is the equation of a linear undamped vibration. The analysis of the numerical solution of this equation could be done quite completely. To

begin with, we make two first-order equations:

$$\frac{dy}{dt} = v \; ; \quad \frac{dv}{dt} = -y$$

These are special cases of the general type $\frac{dy}{dt} = f(y,v)$; $\frac{dv}{dt} = g(y,v)$ which the method described would handle. The functions to be extrapolated are thus y and v themselves.

A. - A complete discussion of the solution will be given for the case of linear extrapolation and integration.

1) Extrapolation and Integration

Suppose that y and v have been evaluated for $t = 0$, h, $2h$, ...$(n-1)h$. We must extrapolate y and v using y_{n-1} and v_{n-1}. The slopes are v_{n-1} and $-y_{n-1}$ respectively, so the extrapolated lines are:

$$y = y_{n-1} + v_{n-1}\left[t - (n-1)h\right]$$

$$v = v_{n-1} - y_{n-1}\left[t - (n-1)h\right]$$

Integrating from $t = (n-1)h$ to $t = nh$

$$\int_{(n-1)h}^{nh} y\,dt = hy_{n-1} + \frac{h^2}{2} v_{n-1}$$

$$\int_{(n-1)h}^{nh} v\,dt = hv_{n-1} - \frac{h^2}{2} y_{n-1}$$

These integrals are combined with y_{n-1} and v_{n-1} to obtain y_n and v_n

$$y_n = y_{n-1} + \int_{(n-1)h}^{nh} v\,dt = y_{n-1}\left(1 - \frac{h^2}{2}\right) + hv_{n-1}$$

$$v_n = v_{n-1} - \int_{(n-1)h}^{nh} y\,dt = v_{n-1}\left(1 - \frac{h^2}{2}\right) - hy_{n-1}$$

These last are the formulas for proceeding from the $(n-1)^{st}$ step to the nth step. It is possible to obtain y_n and v_n as functions of n and thus to examine the long-run behavior of the step-by-step process.

2) Solution of the Difference Equations

An important step in the analysis of a numerical method of the type under discussion is the solution of difference equations. An explanation of the necessary aspects of difference equation theory will be given. A difference equation with one unknown has the form

$$f(y_n, y_{n+1}, \ldots y_{n+p}, n) = 0$$

The type we shall consider will be linear with constant coefficients and of the form

$$A_p y_{n+p} + A_{p-1} y_{n+p-1} + \ldots + A_0 y_n = 0$$

The method of solution is exactly analogous to the solution of linear differential equations with constant coefficients. Let the operator E be that of increasing the index by 1, i.e.

$$E y_n = y_{n+1}.$$

then, the equation becomes

$$(A_p E^p + A_{p-1} E^{p-1} + \ldots + A_0) y_n = 0$$

which, as in the differential equation case, may be factored into

$$A_p (E-r_1)(E-r_2) \ldots (E-r_p) y_n = 0$$

The solution of the simple equation

$$(E-r) y_n = 0 \quad \text{or} \quad y_{n+1} = r y_n$$

is seen to be $y_n = cr^n$ where c is a constant. The complete solution of the difference equation is then

$$y_n = c_1 r_1^n + c_2 r_2^n + \ldots + c_p r_p^n .$$

For example, consider the equation

$$y_{n+2} - 3y_{n-1} + 2y_n = 0$$

and suppose $y_0 = 4$, $y_1 = 7$. Using the equation directly, it is found that

$$y_2 = 13,\ y_3 = 25,\ y_4 = 43 \ldots$$

If we solve the equation $(E^2 - 3E + 2)\, y_n = 0$, the roots are $E = 1$ and $E = 2$, so that $y_n = c_1 \cdot 1^n + c_2 \cdot 2^n$. The initial conditions $y_0 = 4$, $y_1 = 7$, give $c_1 + c_2 = 4$; $c_1 + 2c_2 = 7$, so that $c_1 = 1$, $c_2 = 3$, and $y_n = 1 + 3 \cdot 2^n$.

In case the roots are not distinct, or are complex, modifications analogous to the differential equation case must be made. Since the case of complex roots is important, let us consider such a case. Let,

$$y_{n+2} + y_{n+1} + y_n = 0, \text{ and let } y_0 = 1,\ y_1 = 2.$$

$$(E^2 + E + 1)y_n = 0;\quad E = -\frac{1}{2} \pm i\,\frac{\sqrt{3}}{2} = \cos 120^\circ \pm i \sin 120^\circ$$

$$y_n = c_1(\cos 120^\circ + i \sin 120^\circ)^n + c_2(\cos 120^\circ - i \sin 120^\circ)^n$$

$$= c_1 (\cos n \cdot 120^\circ + i \sin n \cdot 120^\circ) + c_2(\cos n \cdot 120^\circ - i \sin n \cdot 120^\circ)$$

$$= (c_1 + c_2) \cos (n \cdot 120^\circ) + i\,(c_1 - c_2) \sin (n \cdot 120^\circ)$$

and since y_n is real, we can write

$$y_n = c_3 \cos (n \cdot 120^\circ) + c_4 \sin (n \cdot 120^\circ).$$

Substituting initial conditions:

$$c_3 = 1,\quad -\frac{c_3}{2} + \frac{c_4\sqrt{3}}{2} = 2,\quad c_4 = \frac{5}{\sqrt{3}};\ \text{so that}$$

$$y_n = \cos 120n^\circ + \frac{5}{\sqrt{3}} \sin (120n^\circ)$$

Here, $y_2 = -3$, $y_3 = 1$, $y_4 = 2$, $y_5 = -3$, etc., as can be checked by substitutions in the original equation.

The equations which we must solve are simultaneous linear difference equations with constant coefficients.

$$y_n = \left(1 - \frac{h^2}{2}\right) y_{n-1} + h v_{n-1}$$

$$v_n = \left(1 - \frac{h^2}{2}\right) v_{n-1} - h y_{n-1} .$$

Writing in the operational form

$$\left[E - \left(1 - \frac{h^2}{2}\right)\right] y - h v = 0$$

$$h y + \left[E - \left(1 - \frac{h^2}{2}\right)\right] v = 0.$$

To solve, eliminate, say, v.

$$\left[E - \left(1 - \frac{h^2}{2}\right)\right]^2 y - h \left[E - \left(1 - \frac{h^2}{2}\right)\right] v = 0$$

$$h^2 y + h \left[E - \left(1 - \frac{h^2}{2}\right)\right] v = 0$$

Adding, $$\left\{\left[E - \left(1 - \frac{h^2}{2}\right)\right]^2 + h^2\right\} y = 0$$

The roots are

$$E = \left(1 - \frac{h^2}{2}\right) \pm i\, h = r(\cos\theta \pm i \sin\theta)$$

where $$r = \sqrt{1 + h^4/4} \quad ; \quad \tan\theta = \frac{h}{1 - \frac{h^2}{2}} .$$

Thus, $$y_n = c_1 r^n \cos n\theta + c_2 r^n \sin n\theta.$$

To find v_n, substitute in the equation

$$y_{n+1} = \left(1 - \frac{h^2}{2}\right) y_n + h v_n ; \text{ or}$$

$$h v_n = y_{n+1} - \left(1 - \frac{h^2}{2}\right) y_n .$$

Since $h = r \sin\theta$, $1 - \frac{h^2}{2} = r \cos\theta$, this becomes

$$r \sin\theta\, v_n = r^{n+1} c_1 (\cos n\theta \cos\theta - \sin n\theta \sin\theta) +$$

$$r^{n+1} c_2 (\sin n\theta \cos\theta + \cos n\theta \sin\theta) =$$

$$r^{n+1} \cos\theta\, c_1 \cos n\theta - r^{n+1} \cos\theta\, c_2 \sin n\theta$$

$$v_n = \frac{r^n \cos n\theta}{r \sin \theta}\left\{ r\, c_1 \cos\theta + r\, c_2 \sin\theta - r\, c_1 \cos\theta \right\} +$$

$$\frac{r^n \sin n\theta}{r \sin \theta}\left\{ - r\, c_1 \sin\theta + r\, c_2 \cos\theta - r\, c_2 \cos\theta \right\}$$

$$= r^n c_2 \cos n\theta - r^n c_1 \sin n\theta$$

Thus, $y_n = r^n(c_1 \cos n\theta + c_2 \sin n\theta)$

$$v_n = r^n(c_2 \cos n\theta - c_1 \sin n\theta).$$

Now, to satisfy the initial condition, y_0 and v_0

$$y_0 = c_1, \quad v_0 = c_2 .$$

so that finally,

$$y_n = r^n(y_0 \cos n\theta + v_0 \sin n\theta)$$

$$v_n = r^n(v_0 \cos n\theta - y_0 \sin n\theta)$$

where, $r = \sqrt{1 + \frac{h^4}{4}}$ and $\tan\theta = \frac{h}{1 - \frac{h^2}{2}}$.

With these formulas, it is possible to analyze the long-time behavior of the numerical solutions. The points all lie on the curves

$$y = Ae^{\beta t} \sin(\omega t + \alpha)$$

$$v = Ae^{\beta t} \cos(\omega t + \alpha)$$

where, $A = \sqrt{y_0^2 + v_0^2}$, $\beta = \frac{\log r}{h}$, $\omega = \frac{\theta}{h}$ and $\tan\alpha = \frac{y_0}{v_0}$.

It is well to compare the approximations found with the true solutions. They are

$$y = A \sin(t + \alpha)$$

$$v = A \cos(t + \alpha)$$

where A and α are as above.

It can be seen that the points of the approximate solution lie on a damped sine wave. If r exceeds 1, the damping is negative, while, if r is less than 1, the damping is positive. The apparent frequency of the approximating point, ω, is in general different from one.

In the case at hand, r is always greater than 1, and ω is greater than one for small values of h. Therefore, the approximating points will exhibit a sinusoidal behavior which differs from that of the true solution by having a shorter period and an increasing amplitude. For values of h in the vicinity of 3, the negative damping is at its worst. Amplitude rises to millions of times its proper value in only five cycles. If the original differential equation had had a natural frequency different from one, the worst value of h would have been about half the period of the vibration.

E. - Other Extrapolation and Integration Methods

The results of other methods of extrapolation and integration are similar to those just described. The approximating points lie on curves which have the form

$$y = e^{\beta_1 t} \sin(\omega_1 t + \alpha_1) + e^{\beta_2 t} \sin(\omega_2 t + \alpha_2) + \ldots$$

where the values of β, ω, and α are functions of the step length h. One of the β's is approximately zero while the rest are large and negative. The best results were obtained with a parabolic extrapolation and integration using two points and one slope. This resulted in two β's and two ω's. One β was near zero with the corresponding ω very near to one. For small values of h the β was small and negative. For $h = \frac{\pi}{10}$, twenty points per period, only 1% loss in amplitude per period was observed. For values of h near 3, however, the β had a positive maximum so that again negative damping was introduced. In fact, in any numerical method of the type under discussion, negative damping is inevitably introduced for values of h which are large enough.

The attached figures show the behavior of approximation for some of the methods used. Sheet 1 shows the numerical solution for the case of two-point one-slope parabolic extrapolation and integration. There are twenty points per period. Sheets 2 and 3 show the numerical solution for the case ofone-point, one-slope linear extrapolation and integration for two, four, six, eight, and ten points per period.

G. - To generalize the results obtained, the method was applied to the linear vibration equation with damping.

$$\ddot{y} + 2\zeta\dot{y} + y = 0$$

The results can be summarized as follows:

1) Qualitative character of damping was preserved. The only method used was that of linear one-point, one-slop extrapolation and integration. The numerical solution predicted overdamping, critical damping, and underdamping correctly.

2) The amount of damping was usually too small. It was nearly correct for small values of h, but as h increased in size, the amount of damping became much less than the equation required. In fact, for h equal to approximately 2, the negative damping characteristic of the numerical method was enough to cancel all the damping brought in by the equation so that a net negative damping was present.

3) The attached figure shows the apparent damping ratio as a function of number of points per period for different values of true damping ratio. It should be noted that the damping ratio always becomes negative for less than three points per period.

D. - Another generalization was made by investigating the solution of an equation with two natural frequencies

$$\dddot{y} + 101\,\ddot{y} + 100\,y = 0$$

As has been shown, if twenty points per period of an oscillation are taken, the solution is well represented, while, if but two points per period are taken, there is bad divergence. The two natural periods of the above equation are 2π and $.2\pi$, so that when $h = .1\pi$, we have twenty points in the longer period and two in the shorter. Initial conditions were selected to suppress the higher frequency completely. When the numerical solution was carried through, it was found that the higher frequency was not suppressed and the divergence phenomenon associated with two pointsper period was very much in evidence. It appears that only very special initial conditions could suppress the rapid divergence.

PART III - Conclusion

From what has been found it appears that indiscriminate use of numerical methods is not possible. An accurate long-time solution can be found if the steps can be made sufficiently small, or what is the same thing, if the time scale can be expanded sufficiently. The presence of high natural frequencies produces serious distortion if unlimited expansion of time is not possible.

It should be mentioned that in the case studied, the exact solution is known so that the amount by which the numerical solution is in error can be easily seen. Thus, there is no temptation to regard a given numerical

solution as even a close approximation. There is this temptation if the exact solution is not known. If a numerical solution and no other is at hand, one is tempted to regard it as a reasonable approximation to the true solution although it may be as badly in error as were some of the solutions studied above. For this reason it is necessary to confine investigation of a method to equations where an exact solution can be found by other means so that a proper evaluation of the numerical method may be used.

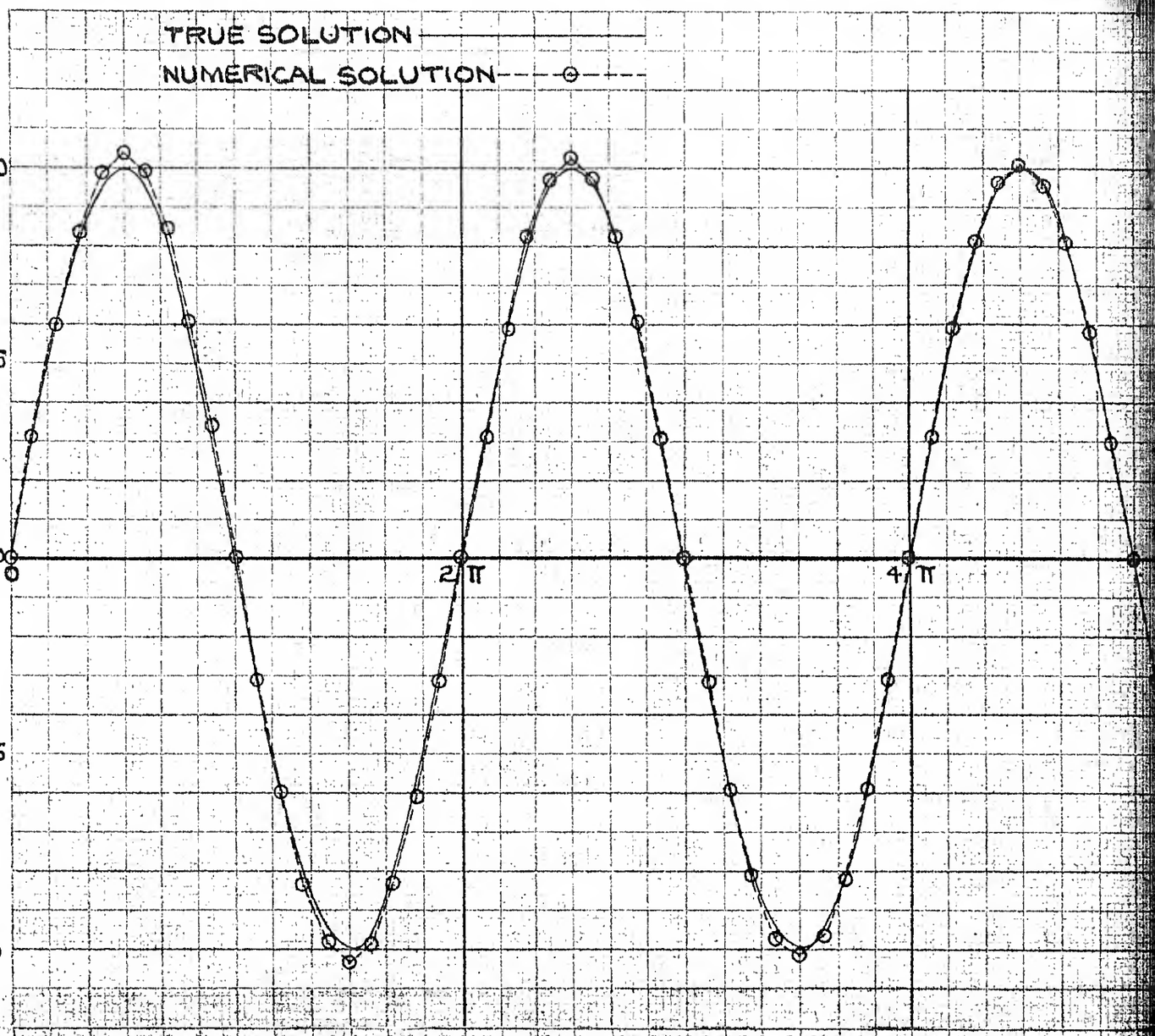
The study made here has taken no account of round-off error. No statment can be made as to the effect of round-off error except that it is greater for smaller time divisions.

It appears that the numerical solution of differential equations where the time scale cannot be indefinitely expanded will have to be effected by more complicated methods. Such methods are at present under consideration.

Engineer: Warren S. Loud

Approved:

WSL:hes



NUMERICAL SOLUTION OF DIFFEREN

EQUATION TO BE SOLVE

INTERPOLATION METHOD USED:

EXTRAPOLATION: PARABOLIC, U

INTEGRATION: PARABOLIC, USI

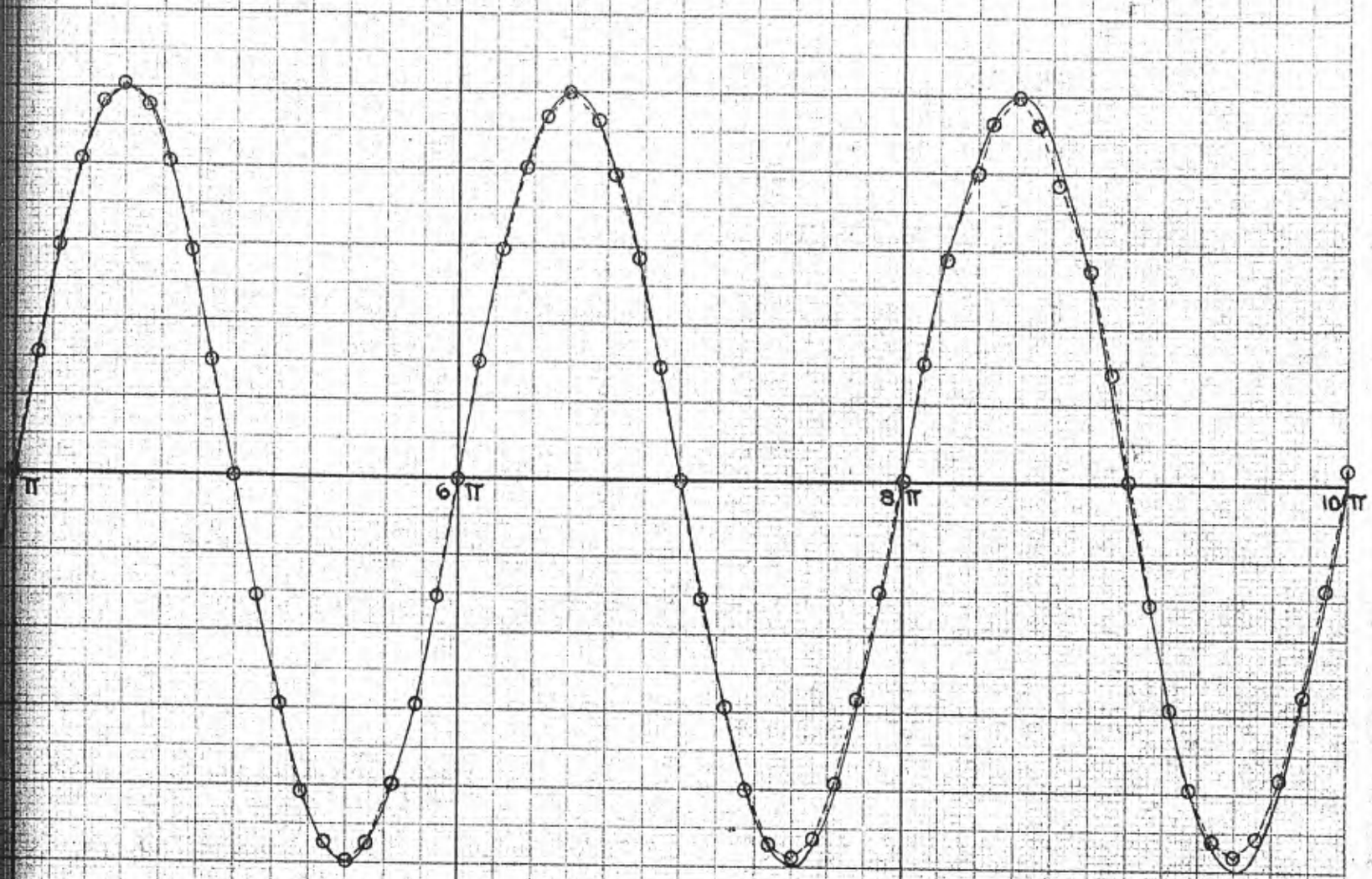
DATA FROM 3MIF122-129

INITI

TRUE SOLUTION OF $\frac{d^2y}{dt^2} = -y$: $y = 1 \sin 1t$

NUMERICAL SOLUTION: $y = e^{-0.002563t}(1.045 \sin 1.0012t - 0.013625$

USED IN 63



OF DIFFERENTIAL EQUATIONS, SHEET 1

TO BE SOLVED: $\frac{d^2y}{dt^2} = -y$

OD USED:

PARABOLIC, USING 2 POINTS AND 1 SLOPE

ARABOLIC, USING 2 POINTS AND 1 SLOPE

-129 20 POINTS PER PERIOD

INITIAL CONDITIONS: $y_0=0$, $\dot{y}_0=1$

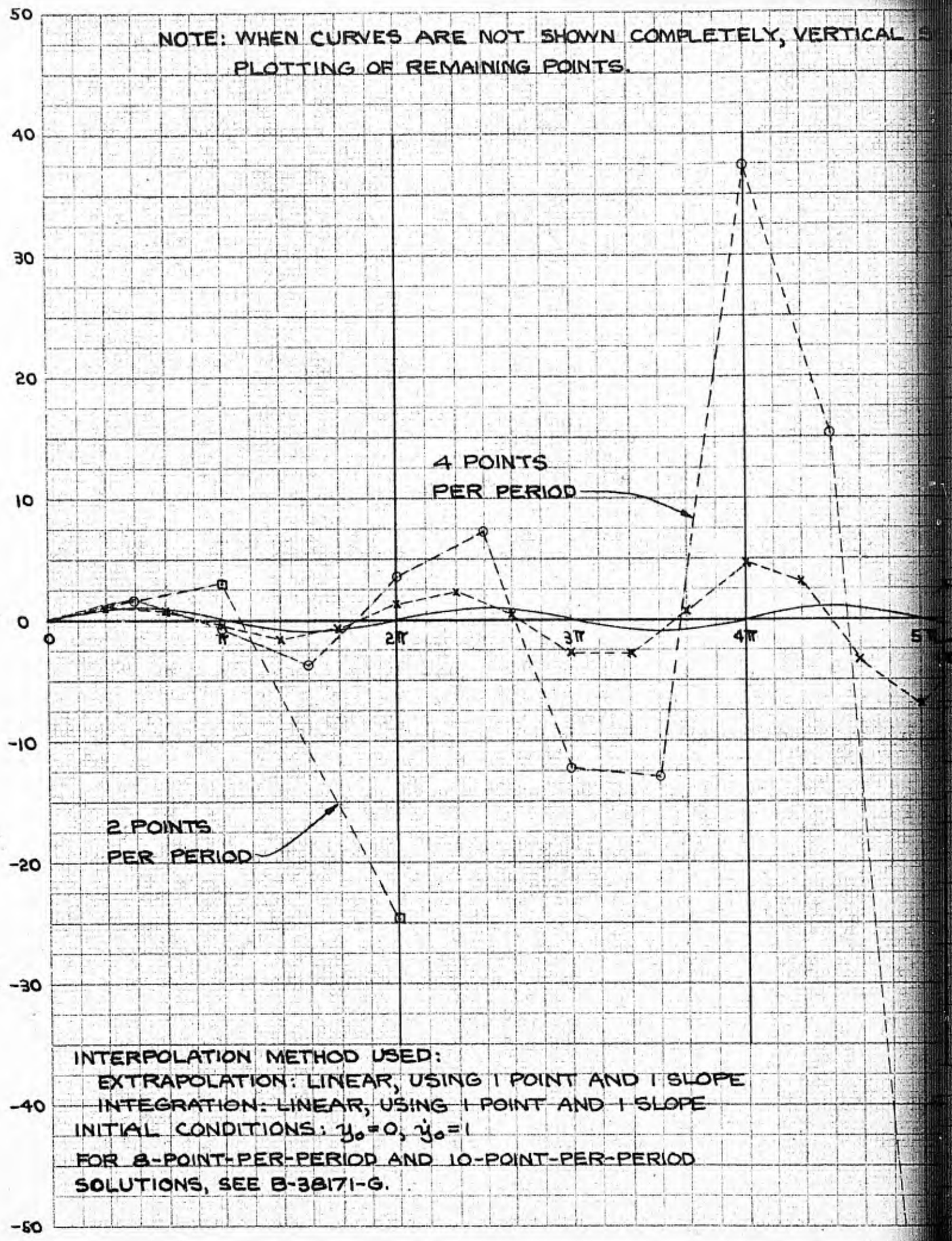
$$2t - 0.013625 \cos 1.0012t) + e^{-7.211t}(0.045 \sin 6t + 0.013625 \cos 6t)$$

REDRAWN 1/21/47

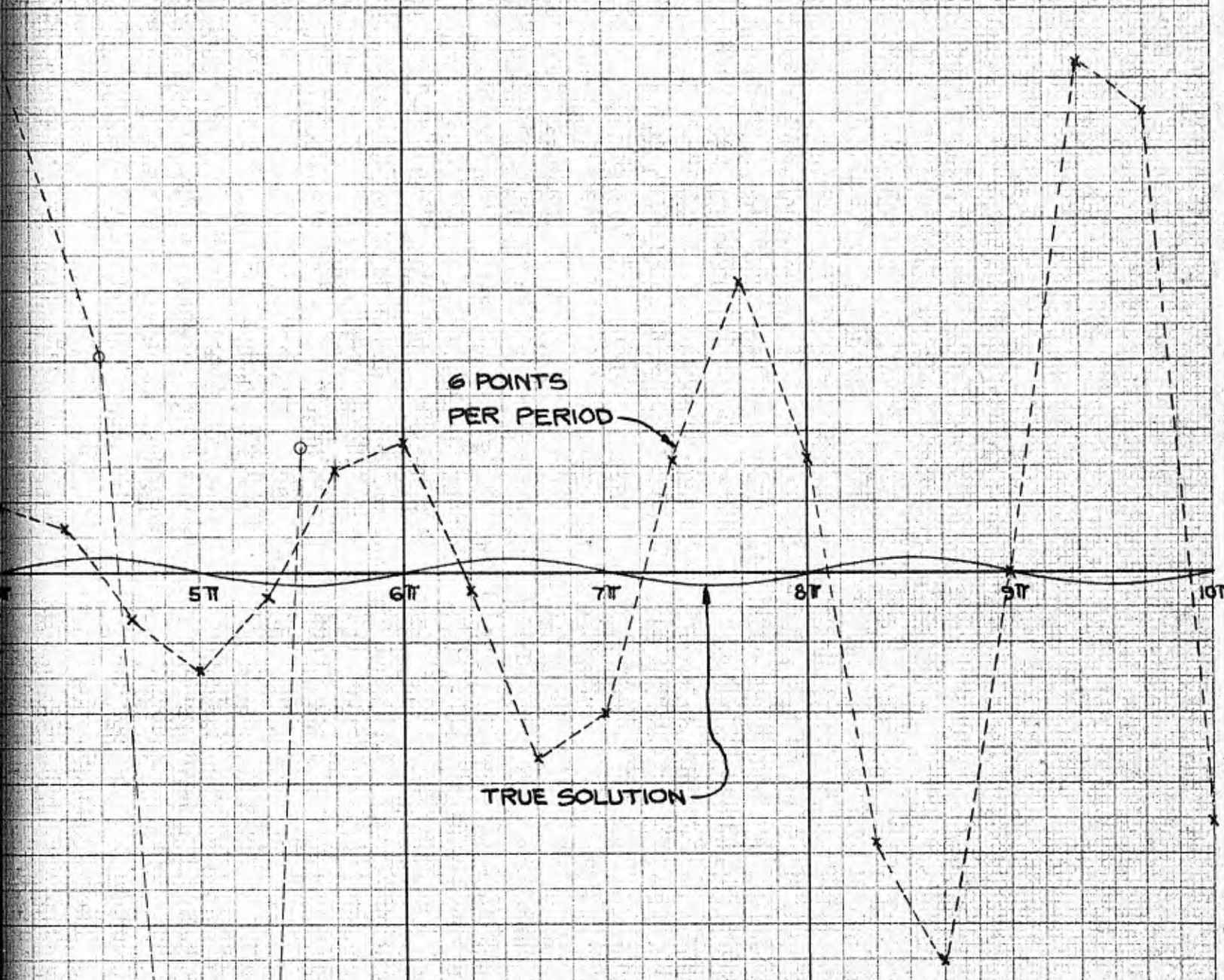
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
SERVOMECHANISMS LABORATORY

S. I. C. NO.	DR D.L.G.	K TL
6345	1/21/47	1/21/47
ENG. WL	APP.	KEL B-38:46-G-1

USED IN 6345 REPORT R-114


NOTE: WHEN CURVES ARE NOT SHOWN COMPLETELY, VERTICAL S
PLOTTING OF REMAINING POINTS.
50
40
30
20
10
0
-10
-20
-30
-40
-50
0
π
2π
3π
4π
5π
4 POINTS
PER PERIOD
2 POINTS
PER PERIOD
INTERPOLATION METHOD USED:
EXTRAPOLATION: LINEAR, USING 1 POINT AND 1 SLOPE
INTEGRATION: LINEAR, USING 1 POINT AND 1 SLOPE
INITIAL CONDITIONS: $y_0 = 0$, $\dot{y}_0 = 1$
FOR 8-POINT-PER-PERIOD AND 10-POINT-PER-PERIOD
SOLUTIONS, SEE B-38171-G.

Y, VERTICAL SCALE IS TOO SMALL TO ALLOW



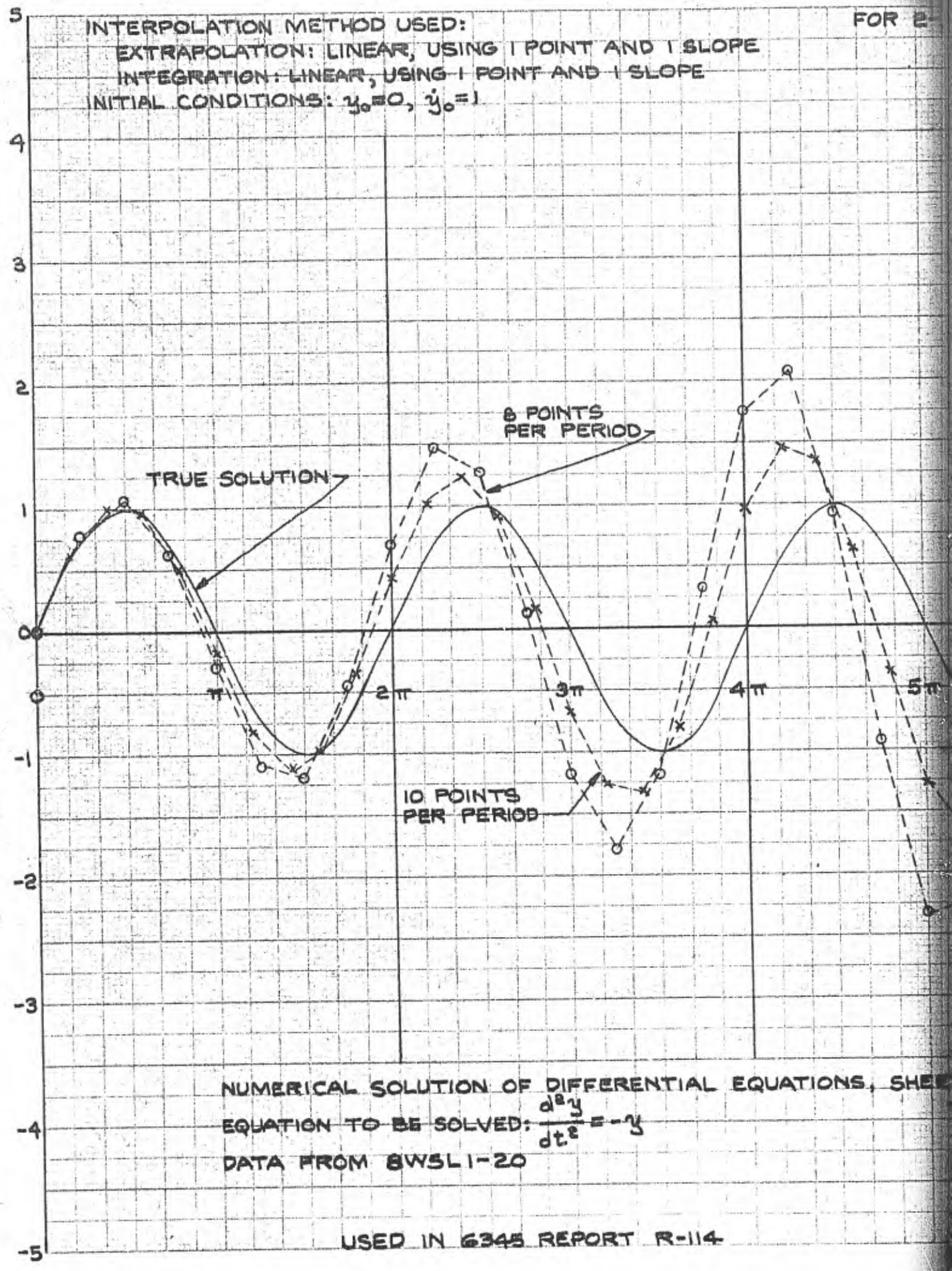
NUMERICAL SOLUTION OF DIFFERENTIAL EQUATIONS, SHEET 2

EQUATION TO BE SOLVED: $\frac{d^2y}{dt^2} = -y$

DATA FROM 8WSL1-20

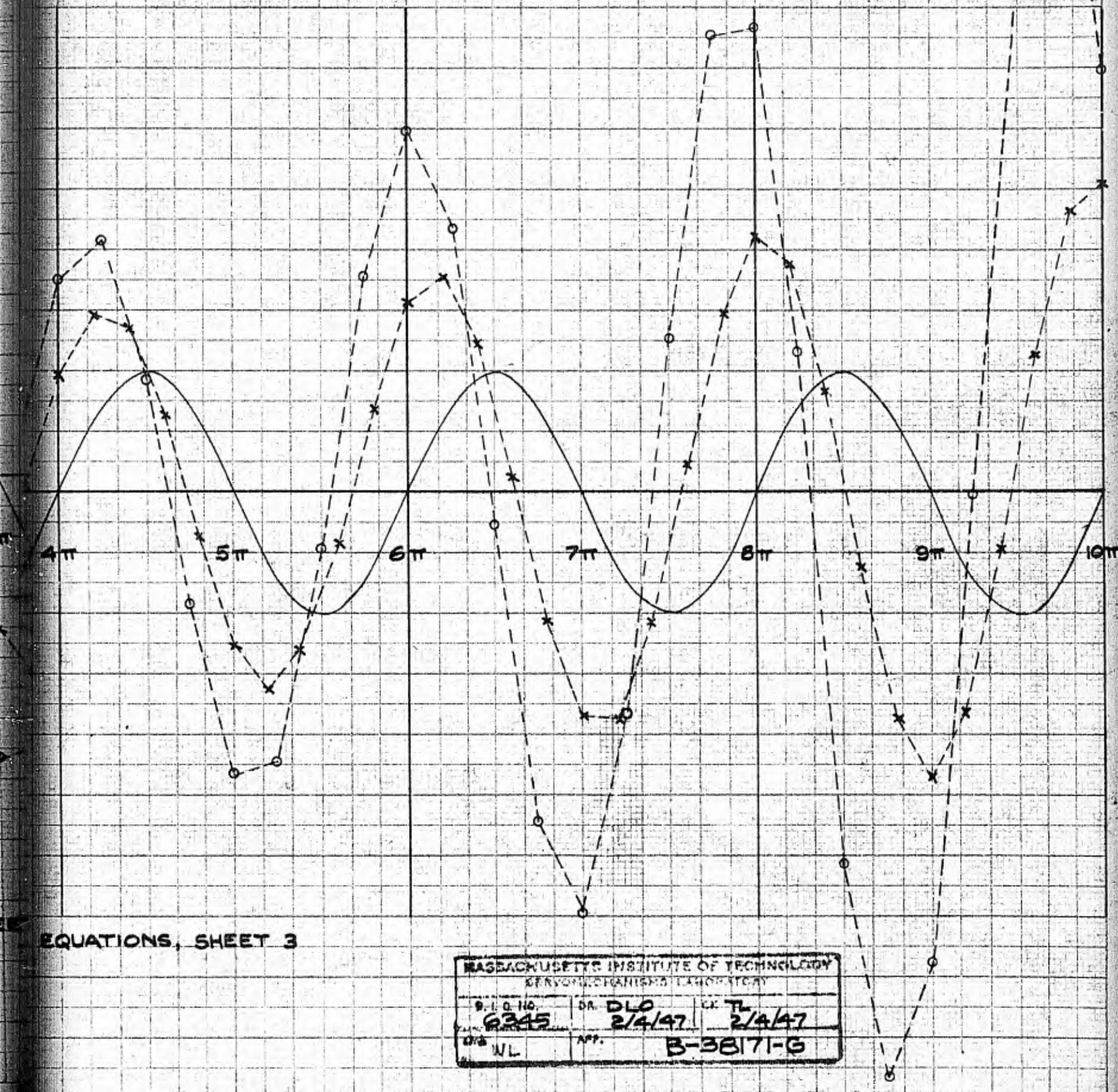
USED IN 6345 REPORT R-114

MASSACHUSETTS INSTITUTE OF TECHNOLOGY SERVOMECHANISMS LABORATORY		
D. I. C. NO. 6345	DR. DLO 1/23/47	K TL 1/23/47
ENG. WL	APP.	B-38170-G


INTERPOLATION METHOD USED:
EXTRAPOLATION: LINEAR, USING 1 POINT AND 1 SLOPE
INTEGRATION: LINEAR, USING 1 POINT AND 1 SLOPE
INITIAL CONDITIONS: $y_0=0$, $\dot{y}_0=1$
FOR 2-
8 POINTS PER PERIOD
TRUE SOLUTION
10 POINTS PER PERIOD
5
4
3
2
1
0
-1
-2
-3
-4
-5
π
2π
3π
4π
5π
NUMERICAL SOLUTION OF DIFFERENTIAL EQUATIONS, SHEE
EQUATION TO BE SOLVED: $\frac{d^2y}{dt^2}=-y$
DATA FROM 8WSL1-20
USED IN 6345 REPORT R-114

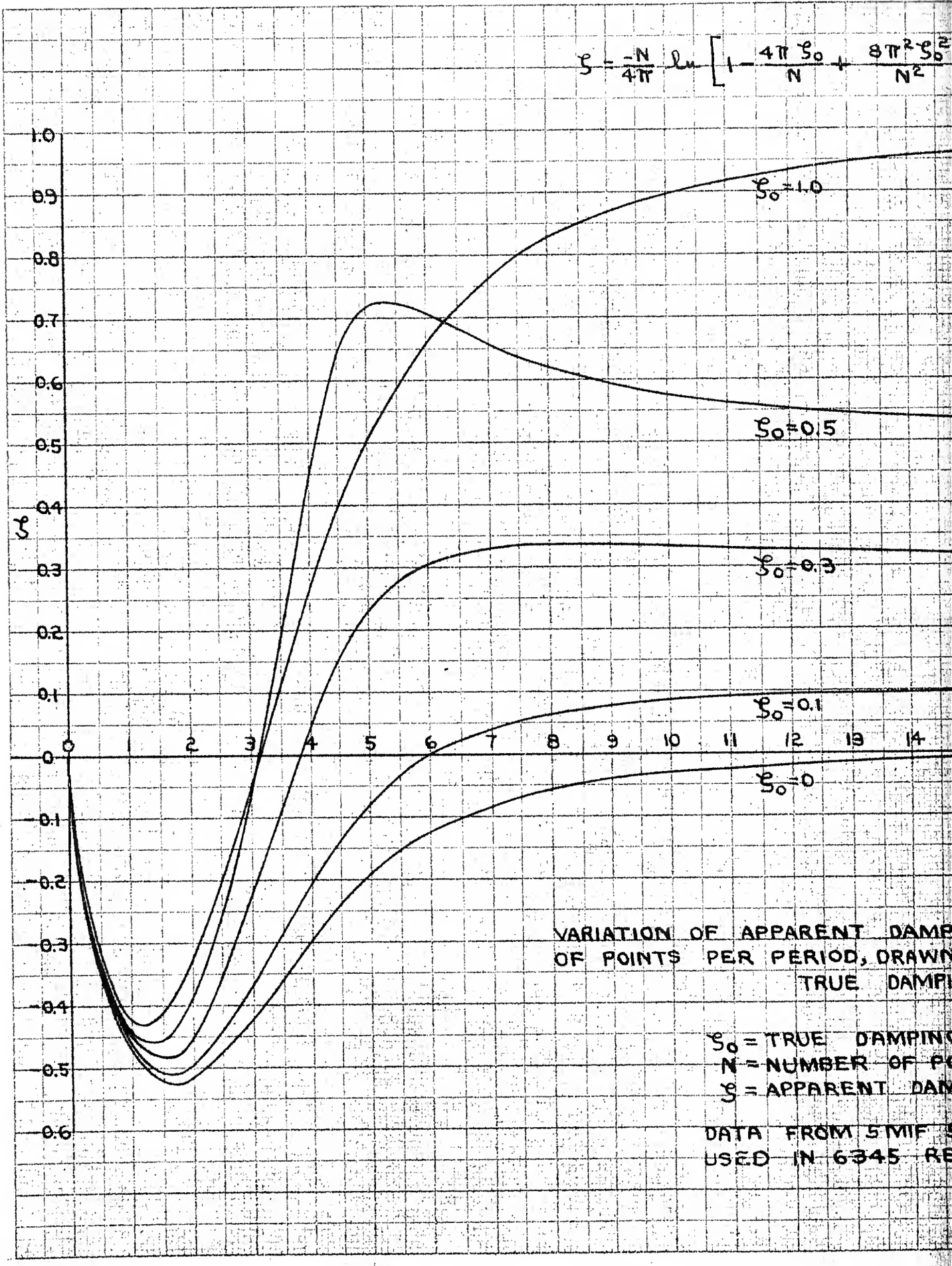
FOR 2-, 4-, AND 6-POINT-PER-PERIOD SOLUTIONS, SEE B-38170-G

PE

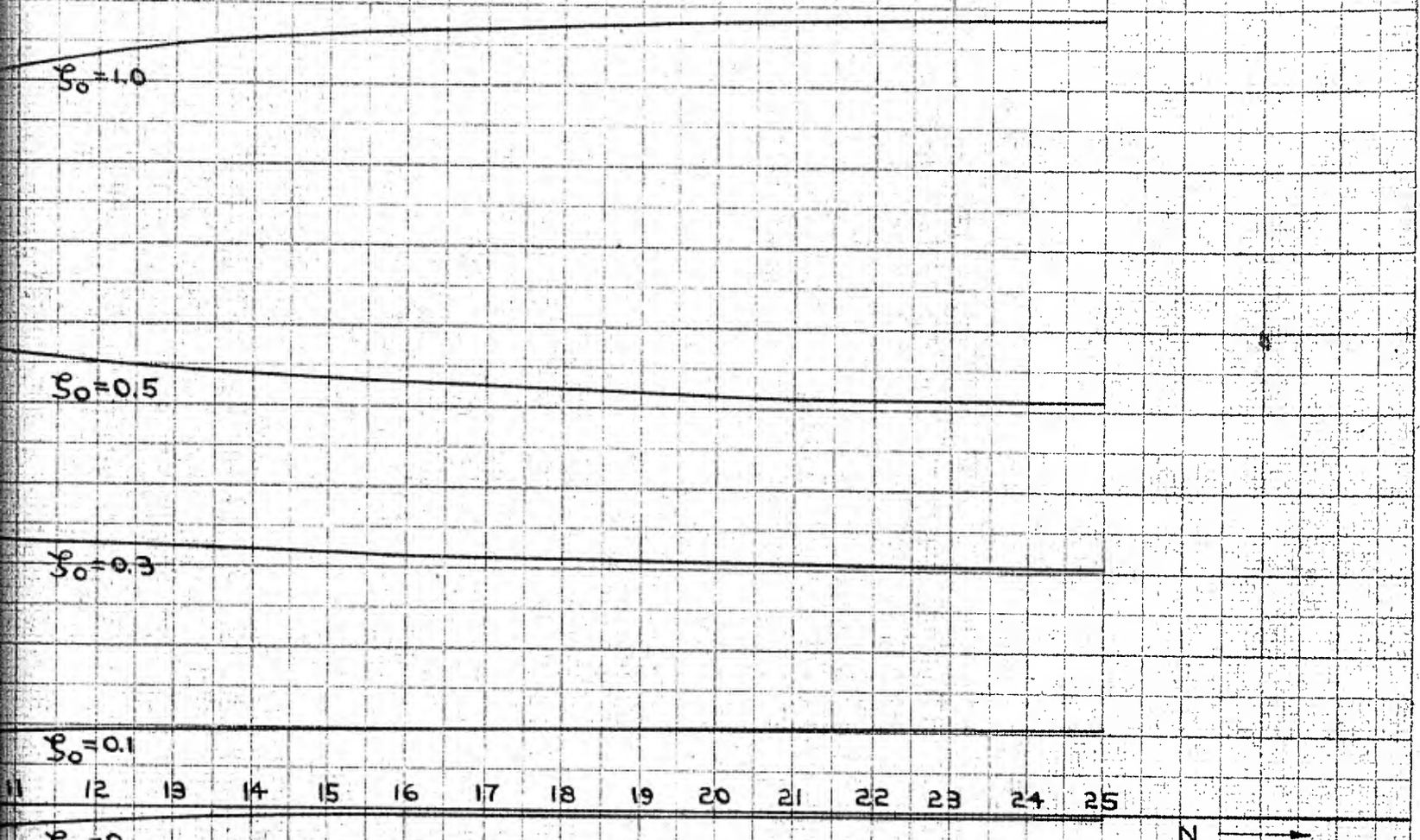


EQUATIONS, SHEET 3

MASSACHUSETTS INSTITUTE OF TECHNOLOGY SERVOMECHANISMS LABORATORY		
D. I. O. No. 6345	DR. DLO 2/4/47	CK. TL 2/4/47
Orig. WL	APP. B-38171-G	


$\zeta = \frac{-N}{4\pi} \ln\left[1 - \frac{4\pi\zeta_0}{N} + \frac{8\pi^2\zeta_0^2}{N^2}\right.$
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
ζ
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
$\zeta_0 = 1.0$
$\zeta_0 = 0.5$
$\zeta_0 = 0.3$
$\zeta_0 = 0.1$
$\zeta_0 = 0$
VARIATION OF APPARENT DAMP
OF POINTS PER PERIOD, DRAWN
TRUE DAMPI
ζ_0 = TRUE DAMPIN
N = NUMBER OF P
ζ = APPARENT DAM
DATA FROM SMIF
USED IN 6345 RE

$$\left[1 - \frac{4\pi\zeta_0}{N} + \frac{8\pi^2\zeta_0^2}{N^2} - \frac{8\pi^3\zeta_0}{N^3} + \frac{4\pi^4}{N^4}\right]$$



APPARENT DAMPING RATIO WITH NUMBER
ER PERIOD, DRAWN FOR SEVERAL VALUES OF
TRUE DAMPING RATIO.

= TRUE DAMPING RATIO
= NUMBER OF POINTS PER PERIOD
= APPARENT DAMPING RATIO

TA FROM 5 MIF 58-63
ED IN 6345 REPORT R-114

EXTRAPOLATION AND INTEGRATION:
LINEAR, 1 POINT AND 1 SLOPE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
SERVOMECHANISMS LABORATORY
B. NO. 6345 DR PLO 2/21/47 CK TL 2/21/47
SHT. APP. WL
B-38174-G-1

KEUFFEL & ESSER CO. N. Y.

Memorandum No. M-63

To: J. W. Forrester

From: Warren S. Loud

Subject: Discussion of Numerical Methods

Date: March 17, 1947

Digital Computation and Numerical Methods as a New Field of Mathematics

With the advent of high-speed digital computing equipment, the need for more study of numerical methods becomes apparent. Many questions of a mathematical nature arise when an attempt is made to employ this new machinery.

Two main aspects of the theory present themselves. First there is the question of errors. A numerical solution of an analytic problem will be in error because the numerical method is discrete and not continuous, and because any numerical expression is only an approximation to a quantity which is not rational with a small denominator. The two kinds of errors are called truncation errors and round-off errors respectively. The study of errors is of utmost importance. No confidence can be place in a numerical result unless an estimate of its error is available. A numerical solution of a process can be badly in error yet no clue to its error might be in evidence.

A second aspect of the theory is the question of numerical methods used with physical problems without an intervening differential equation. It would be of interest to know just how numerical methods can be used directly with physical or engineering problems.

Ideas for a Course in Numerical Methods

I. Background

At first sight numerical methods appear to be merely a tool for understanding and treatment of mathematical subjects. This is not entirely true, for numerical computation and its accompanying problems form a proper field of pure mathematics. Indeed one can look at the problem from two points of view, that of using numerical methods to learn more about other fields of mathematics or that of using the other fields of mathematics to learn more about numerical methods. However, since numerical methods do have much application to other branches of mathematics and also to engineering, it is advisable to obtain as broad a background as possible in the application of numerical methods to these other fields. There is a three-fold objective here. First, it is well to know the material so that places where numerical methods are of use at present can be known. Second, a wide acquaintance with related fields is desirable if new results within them are to be found by numerical methods. Third, it is well to have as many tools as possible for advancing the art of computational methods.

II. Specific Topics

Mathematical fields which would be studied include:

A. Algebraic and transcendental equations.

B. Differential Equations both ordinary and partial.

C. Integral Equations.

D. General Computations.

Discussion

A. Under algebraic equations we might emphasize the field of simultaneous linear equations in large numbers of unknowns. In general, iterative procedures should be emphasized in this field in view of the nature of high-speed computers currently being developed.

B. Ordinary Differential Equations

Numerical methods to date have not been helpful in the theory of differential equations because the time involved to obtain sufficient information to generalize from particular solutions has been prohibitive. With the new machines it is to be anticipated that sufficient particular information can be gathered in a reasonable time so that generalizations can be drawn. This can be considered a reason for study of this field over and above the desire to solve equations which cannot be solved in terms of elementary functions.

Partial Differential Equations

The statements made under ordinary differential equations apply here but with more force. Partial differential equations are more difficult to solve. Nevertheless many physical problems are currently phrased in terms of partial differential equations. Also since the theory is so much more limited, it will be important to have means for gaining new knowledge.

C. Integral Equations

Integral equations have been used more in recent years in physical problems. It would seem that here is a place where numerical methods could be widely used.

D. General Computations

There are problems in existence today which do not yield to analytical solution. Such problems arise in the theory of design of optical systems. Numerical methods are absolutely necessary in such problems. In addition matrix calculations can be studied from the point of view of high-speed machines.

III. Mathematical Methods Used

A. Relaxation Methods

B. Iteration Methods

C. General Numerical Methods

Relaxationsmethods are a recently developed tool finding much use in numerical work with engineering problems.

Interation procedures are given a special place because it appears that they will be of highest importance with high-speed machines. They represent an essentially simple process repeated many times.

In addition general numerical methods as described in books like Scarborough would be considered.

Under numerical methods in general, two remarks should be made. First, there needs to be an assembly of pertinent material. Much valuable literature on the subject is not collected in books, and it is often not available in English. Second, numerical methods should be examined with the thought in mind that they are to be used with high-speed machines. Present-day methods are designed for desk calculator speeds or slower. As a consequence, they are complicated, calling for reasoning and judgment on the part of the

operator. For high-speed machines, simpler repetitive methods should be considered. The high-speeds permit so many steps to be taken that the accuracy is still increased.

IV. Errors

Numerical methods introduce errors. Such errors are important, since it is necessary to know how close to the true solution a numerical solution lies. They fall in general into two classes, truncation errors and round-off errors. Truncation errors are those due to the fact that a numerical process is different from an analytical process. Such errors can be studied by analytical methods. Round-off errors are due to the necessary rounding off of each numerical step. This introduces errors which are of a random nature and difficult to analyze.

New Mathematical Results which can be Obtained by Numerical Methods

At the present time it is difficult to state just what new mathematical results can be obtained with high-speed machines. An opportunity for mathematical experiment presents itself. The technique of generalizations from particular experimental results so long used in science can be greatly extended in mathematics. Much work of this nature ought to be done.

As an example of results which can be obtained the theory of partial differential equations is in need of particular results that more generalizations may be drawn. There will also be an opportunity to study long-time iteration procedures in general. Iteration procedures have not been studied much in the past because they converge slowly if at all. However with high-speed

machines the slowness of convergence is not of importance and the properties of such procedures can be studied.

Many theorems in mathematics are of such a general nature that their use in specific cases is not possible. With machines of the type considered more concrete information can be obtained from these general theorems. For example, an existence theorem in order to be general often gives too restricted a result. It should be possible to obtain more comprehensive results in particular cases with such methods.

Warren S. Lund

Warren S. Lund

vh-

MEMORANDUM NO. M-73*

SERVOMECHANISMS LABORATORY
Massachusetts Institute of Technology
Cambridge, Massachusetts

TO: Jay W. Forrester 6345

FROM: Warren S. Loud Page 1 of 4 pages

SUBJECT: Discussion of Determinants of Large Order by Professor Norbert Wiener

DATE: May 1, 1947

Computation of determinants of high order involves problems which do not appear for low orders. The error present in the terms is always multiplied by a factor proportional to the square root of the order. This results in the situation that for large determinants, if the terms are not known accurately, the value found is quite meaningless. Standard computation methods introduce more error than the above inevitable error because of subtractions of large numbers.

Professor Wiener gave a description of these two errors, the unavoidable and the avoidable, with suggestions for reducing the latter. He then proceeded to discuss places in mathematics where determinants are important, in particular, integral equations.

I. Mean Square Size of Large Determinants

Let the elements of a random determinant be a_{ij}. $i,j = 1,2,\ldots,n$. Let the mean of each element be zero, and let the r.m.s. of each element be a. Also assume the elements are completely independent.

Then, if D is the determinant of the a_{ij}, D^2 is composed of n! squares of terms each of which has a mean of a^{2n}, and the cross terms of mean zero. Thus, the r.m.s. of the determinant is $\sqrt{n!}\ a^n$.

II. Effect of Errors in the Terms

Suppose there is an error b_{ij} in each term where the b_{ij}'s have mean zero, r.m.s. b, and are independent of each other and of the a_{ij}. We seek the r.m.s. error.

If D' is the determinant of $a_{ij} + b_{ij}$, the above says the r.m.s. of D' is $\sqrt{n!}\ (a^2 + b^2)^{n/2}$.

* Revised November 17, 1947

Now $D' - D$ contains all the terms of D' which have at least one b in them. Therefore $(D' - D)^2$ contains all the square of these terms plus terms of zero mean. The mean of $(D' - D)^2$ will be $\left[(a^2 + b^2)^n - a^{2n}\right] n!$ and the r.m.s. of $D' - D$ will be $\sqrt{n!}\ \sqrt{(a^2 + b^2)^n - a^{2n}}$

The relative error is

$$\frac{\sqrt{n!}\ \sqrt{(a^2 + b^2)^n - a^{2n}}}{\sqrt{n!}\ a^n} = \sqrt{\left(1 + \frac{b^2}{a^2}\right)^n - 1}$$

III. Largest Tolerable Error in a_{ij}

An error in a determinant of over 100% is intolerable. Since an important property of a determinant is its being different from zero or being zero, a 100% error introduces uncertainty here and is intolerable. Therefore, we must have

$$\sqrt{\left(1 + \frac{b^2}{a^2}\right)^n - 1} \leq 1$$

$$1 + \frac{b^2}{a^2} \leq 2^{\frac{1}{n}}$$

$$\frac{b^2}{a^2} \leq \left(2^{\frac{1}{n}} - 1\right)$$

Now, if n is large, $2^{\frac{1}{n}} - 1 = \frac{\log 2}{n}$ approximately

so, $\frac{b^2}{a^2} \leq \frac{\log 2}{n}$ $\quad \left|\frac{b}{a}\right| \leq \sqrt{\frac{\log 2}{n}} = \frac{.833}{\sqrt{n}}$.

Conversely, if $\left|\frac{b}{a}\right| = \epsilon$, the r.m.s. relative error is $\sqrt{(1 + \epsilon^2)^n - 1} = \sqrt{n}\,\epsilon$ approximately, so that all errors in terms are multiplied by $\sqrt{n}$.

If n = 100, we must have terms to within .083 and small errors are multiplied approximately by 10.

IV. Discussion of Methods of Evaluation

If a determinant is evaluated by evaluating the n! terms, we find we obtain numbers of the order of $\sqrt{n!}\ a^n$ by subtracting numbers of the order of $\frac{n!}{2}\ a^n$. If n is large, this is very wasteful of significant figures and may destroy all of them.

For large determinants it is better to obtain relatively small numbers by multiplying rather than subtracting. The method Wiener suggested is the following:

1. Divide the determinant row by row by the main-diagonal elements, reducing them all to unity. The determinant so obtained

$$\begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & 1 \end{vmatrix}$$

is multiplied by

$$\begin{vmatrix} 1 & -\epsilon a_{12} & \cdots & -\epsilon a_{1n} \\ -\epsilon a_{21} & 1 & \cdots & -\epsilon a_{2n} \\ & & \cdots & \\ -\epsilon a_{n1} & -\epsilon a_{n2} & \cdots & 1 \end{vmatrix}$$

The resulting determinant is treated the same way, and the process is iterated until the resulting determinants are as close as desired to the identity. If ϵ is taken too large, the process will not converge. If ϵ is taken too small, the process will converge very slowly.

(The above process was tried for second-order determinants. It can diverge badly, and moreover although for suitable ϵ it will converge, it can converge to a value more than 100% off.)

Professor Wiener went on to remark that we might as well study determinants which differed but little from the identity.

$$D = \begin{vmatrix} 1 + b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & 1 + b_{22} & \cdots & b_{2n} \\ & \cdots & & \\ b_{n1} & b_{n2} & \cdots & 1 + b_{nn} \end{vmatrix}$$

We find $D = 1 + \sum b_{nn} + \frac{1}{2!} \sum \sum \begin{vmatrix} b_{mm} & b_{mn} \\ b_{nm} & b_{nn} \end{vmatrix} + \cdots$

the mean of D^2 is $1 + nb^2 + \frac{n(n-1)}{2!} \cdot 2! \; b^4 + \frac{n(n-1)(n-2)}{3!} \; 3! \; b^6 + \cdots$

If $nb^2 = 1$, $D^2 = 1 + 1 + \left(1 - \frac{1}{n}\right) + \cdots$ and D^2 is nearly n.

If $nb^2 = k$, with $k < 1$ and neglecting $\frac{1}{n}$, $\frac{2}{n}$ ... compared to 1,

$$D^2 < 1 + k + k^2 + \cdots k^n$$

$$< \frac{1}{1 - k}$$ which equals 2 when $k = \frac{1}{2}$.

Thus $b \approx \frac{1}{\sqrt{2n}}$ or $\frac{.7}{\sqrt{n}}$ gives an error less than 100%

but $b \approx \frac{1}{\sqrt{n}}$ gives an error greater than 100%.

For some a, $.7 < a \leq 1$, $b = \frac{a}{\sqrt{n}}$, gives an error of 100%.

The solution of linear algebraic equations can be effected without the use of determinants, since there exist other methods which involve less computation and less chance for error.

The Fredholm theory of integral equations does involve determinants and this is a place where numerical methods with determinants definitely have a place.

Warren S. Loud

WSL/has

Warren S. Loud

MEMORANDUM NO. M-78, Revision 1

SERVOMECHANISMS LABORATORY
Massachusetts Institute of Technology
Cambridge, Massachusetts

TO: Jay W. Forrester, R. R. Everett, F. E. Swain, M. I. Florencourt — 6345

FROM: Warren S. Loud — Page 1 of 6

SUBJECT: Solution of Equations with Minimum Error

DATE: June 2, 1947

The following is a report of the discussion given on May 12, 1947, by Professor Wiener for the representatives of the Departments of Electrical Engineering and Mathematics.

When a numerical process is used, "obvious" methods of setting up a numerical problem often introduce excessive error. Let us consider the problem of finding the derivative of an approximated function. A process that gives an exact derivative is very undesirable. If a function is approximated to a finite number of decimal places, the approximation is a step-function, whose derivative is zero or infinite at all points, so that the exact derivative of the approximation is of no value at all.

Let us try to find a process which will find the derivative of $f(t)$ with the least error in the mean-square sense. Given an approximation to $f(t)$, $f(t) + g(t)$, where $g(t)$ is the error, we seek an operation on $f + g$ which will give $f'(t)$ with least mean-square error. We seek an operator $K(t)$ such that

$$\int_{-\infty}^{t} K(t-\tau) \left[f(\tau) + g(\tau) \right] d\tau$$

is approximately $f'(t)$. Precisely we seek $K(t)$ such that the mean with respect to t of

$$\left[\int_{-\infty}^{t} K(t-\tau) \left[f(\tau) + g(\tau) \right] d\tau - f'(t) \right]^2$$

is a minimum. The mean with respect to t of a function $f(t)$ is defined as

$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\,dt$. We abbreviate this as $\underset{t}{M} f(t)$.

The expression to be minimized is

$$\underset{t}{M} \int_{-\infty}^{t}\int_{-\infty}^{t} K(t-\tau)\left[f(\tau)+g(\tau)\right] K(t-\sigma)\left[f(\sigma)+g(\sigma)\right] d\tau\, d\sigma$$
$$- 2\underset{t}{M} \int_{-\infty}^{t} K(t-\tau)\left[f(\tau)+g(\tau)\right] f'(t)\,d\tau + \underset{t}{M}\left[f'(t)\right]^2.$$

To simplify the above we need some knowledge of $f(t)$ and $g(t)$. Let us assume that we know the auto-correlations of $f(t)$ and $g(t)$ and their cross correlation. These are respectively

$$\phi_{11}(\tau) = \underset{t}{M}\left[f(t+\tau)\, f(t)\right]$$

$$\phi_{22}(\tau) = \underset{t}{M}\left[g(t+\tau)\, g(t)\right]$$

$$\phi_{12}(\tau) = \underset{t}{M}\left[f(t+\tau)\, g(t)\right]$$

Let us first make the change of variable $t-\tau=\tau_1$ $t-\sigma=\sigma_1$

The expression to be minimized becomes

$$\underset{t}{M} \int_{0}^{\infty}\int_{0}^{\infty} K(\tau_1)\,K(\sigma_1)\left[f(t-\tau_1)+g(t-\tau_1)\right]\left[f(t-\sigma_1)+g(t-\sigma_1)\right] d\tau_1\, d\sigma_1$$
$$-2\underset{t}{M}\int_{0}^{\infty} K(\tau_1)\left[f(t-\tau_1)+g(t-\tau_1)\right] f'(t)\,d\tau_1 + \underset{t}{M}\left[f'(t)\right]^2$$
$$= \int_{0}^{\infty}\int_{0}^{\infty} K(\tau)\,K(\sigma)\left[\phi_{11}(\sigma-\tau)+\phi_{12}(\sigma-\tau)+\phi_{12}(\tau-\sigma)+\phi_{22}(\sigma-\tau)\right] d\tau\, d\sigma$$
$$-2\int_{0}^{\infty} K(\tau)\left[\phi_{11}'(\tau)+\phi_{12}'(\tau)\right] d\tau + \underset{t}{M}\left[f'(t)\right]^2$$

where τ_1 and σ_1 have been written as τ and σ. Noting that the third term

is independent of K, we write

$$A = \int_0^\infty \int_0^\infty K(\tau)\, K(\sigma)\, \phi\,(\sigma - \tau)\, d\tau d\sigma - 2 \int_0^\infty K(\tau) \left[\phi_{11}'\,(\tau) + \phi_{12}'\,(\tau) \right] d\tau$$

to be minimized.

To do this, we write $K + \epsilon(\delta K)$ in place of K. The resulting expression should be minimum for $\epsilon = 0$ no matter what δK is.

$$A = \int_0^\infty \int_0^\infty \left[K(\tau) + \epsilon\, \delta K(\tau) \right] \left[K(\sigma) + \epsilon\, \delta K(\sigma) \right] \phi\,(\sigma - \tau)\, d\tau\, d\sigma$$

$$- 2 \int_0^\infty \left[K(\tau) + \epsilon\, \delta K(\tau) \right] \left[\phi_{11}'\,(\tau) + \phi_{12}'\,(\tau) \right] d\tau$$

$$\frac{dA}{d\epsilon} = \int_0^\infty \int_0^\infty \left[K(\tau)\, \delta K(\sigma) + K(\sigma)\, \delta K(\tau) + 2\epsilon\, \delta K(\tau)\, \delta K(\sigma) \right] \phi(\sigma - \tau)\, d\sigma d\tau$$

$$- 2 \int_0^\infty \delta K(\tau) \left[\phi_{11}'\,(\tau) + \phi_{12}'\,(\tau) \right] d\tau$$

For $\epsilon = 0$, this is

$$\frac{dA}{d\epsilon} = \int_0^\infty \int_0^\infty 2K(\sigma)\, \delta K(\tau)\, \phi\,(\sigma - \tau)\, d\sigma\, d\tau - 2 \int_0^\infty \delta K(\tau) \left[\phi_{11}'\,(\tau) + \phi_{12}'\,(\tau) \right] d\tau$$

since $\phi\,(\sigma - \tau) = \phi\,(\tau - \sigma)$. That is,

$$2 \int_0^\infty \delta K(\tau) \left\{ \int_0^\infty K(\sigma)\, \phi\,(\sigma - \tau)\, d\sigma - \phi_{11}'\,(\tau) - \phi_{12}'\,(\tau) \right\} d\tau = 0$$

for all possible $\delta K(\tau)$. This is possible only if

$$\int_0^\infty K(\sigma)\, \phi\,(\sigma - \tau)\, d\sigma = \phi_{11}'\,(\tau) + \phi_{12}'\,(\tau)$$

which is an integral equation which K must satisfy for $\tau \geq 0$.

To solve the integral equation, take Fourier Transforms.

$$\int_{-\infty}^{\infty} \left\{ \int_{0}^{\infty} K(\sigma)\ \phi'(\sigma-\tau)\, d\sigma - \phi_{11}'(\tau) - \phi_{12}'(\tau) \right\} e^{-i\omega\tau}\, d\tau$$

Since the expression in braces is zero for $\tau > 0$, the transform is of the form

$$\int_{-\infty}^{0} G(\tau)\, e^{-i\omega\tau}\, d\tau$$

and so is bounded for Im $\omega > 0$.

We manipulate the transform further and obtain

$$\int_{0}^{\infty} \int_{-\infty}^{\infty} K(\sigma)\, \phi(\tau-\sigma)\, e^{-i\omega\sigma}\, e^{-i\omega(\tau-\sigma)}\, d\sigma d\tau - \int_{-\infty}^{\infty} \left[\phi_{11}'(\tau) + \phi_{12}'(\tau) \right] e^{-i\omega\tau}\, d\tau$$

since $\phi(\tau-\sigma) = \phi(\sigma-\tau)$,

$$= \int_{0}^{\infty} K(\sigma)\, e^{-i\omega\sigma} d\sigma \int_{-\infty}^{\infty} \phi(\tau) e^{-i\omega\tau}\, d\tau - \int_{-\infty}^{\infty} \left[\phi_{11}'(\tau) + \phi_{12}'(\tau) \right] e^{-i\omega\tau}\, d\tau$$

where $\tau - \sigma$ has been replaced by τ in the τ integral.

Now write

$$\int_{0}^{\infty} K(\sigma) e^{-i\omega\sigma}\, d\sigma = k(\omega)$$

$$\int_{-\infty}^{\infty} \phi(\tau) e^{-i\omega\tau}\, d\tau = \overline{\phi}(\omega)$$

$$\int_{-\infty}^{\infty} \left[\phi_{11}'(\tau) + \phi_{12}'(\tau) \right] e^{-i\omega\tau}\, d\tau = H(\omega)$$

We note that $k(\omega)$ is bounded for Im $\omega < 0$.

Assume $\underset{-}{\overset{+}{\Phi}}(\omega)$ is a rational function. Separate it into two factors $\Psi_1(\omega)$ and $\Psi_2(\omega)$ which are free from zeros and poles in the lower and upper half planes respectively.

We thus have,

$k(\omega)\ \underset{-}{\overset{+}{\Phi}}(\omega) - H(\omega)$ bounded for Im $\omega \geq 0$

$$k(\omega)\ \Psi_1(\omega) - \frac{H(\omega)}{\Psi_2(\omega)} \text{ bounded for Im } \omega \geq 0$$

since Ψ_2 has no zeros for Im $\omega \geq 0$.

Now, if

$$c(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{H(\mu)}{\Psi_2(\mu)}\, e^{i\mu\tau}\, d\mu,$$

$$\frac{H(\omega)}{\Psi_2(\omega)} = \int_{-\infty}^{\infty} c(\tau) e^{-i\omega\tau}\, d\tau = \int_{-\infty}^{0} c(\tau) e^{-i\omega\tau}\, d\tau + \int_{0}^{\infty} c(\tau) e^{-i\omega\tau}\, d\tau$$

$$= A(\omega) + B(\omega) \text{ say.}$$

$k(\omega)\ \Psi_1(\omega) - A(\omega) - B(\omega)$ is bounded for Im $\omega \geq 0$

But, $A(\omega)$ is bounded for Im $\omega \geq 0$

therefore, $k(\omega)\ \Psi_1(\omega) - B(\omega)$ is bounded for Im $\omega \geq 0$.

But, $k(\omega)$, $\Psi_1(\omega)$, and $B(\omega)$ are all bounded for Im $\omega \leq 0$.

therefore $k(\omega)\ \Psi_1(\omega) - B(\omega)$ is bounded for all ω and so is a constant which may be taken as zero.

$$\therefore \quad k(\omega) = \frac{1}{\Psi_1(\omega)} B(\omega)$$

$$= \frac{1}{\Psi_1(\omega)} \int_0^\infty c(\tau) e^{-i\omega\tau} \, d\tau$$

$$= \frac{1}{\Psi_1(\omega)} \int_0^\infty \int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{R(\mu)}{\Psi_2(\mu)} e^{i\mu\tau} e^{-i\omega\tau} \, d\mu \, d\tau$$

$$= \frac{1}{2\pi \Psi_1(\omega)} \int_0^\infty \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\phi_{11}^{1}(\sigma) + \phi_{12}^{0}(\sigma)}{\Psi_2(\mu)} e^{-i\mu\sigma} e^{i\mu\tau} e^{-i\omega\tau} \, d\sigma \, d\mu \, d\tau$$

In conclusion we mention certain changes that may occur in non-linear cases. Message and noise may no longer be additive. The best approximation may no longer be mean-square. We should try to use only the past of the messages. For almost all cases, the future distribution can be obtained. Finally, we emphasize that the original set-up of a problem will not always be the proper set-up for minimum error.

Warren S. Loud

Warren S. Loud

c. Prof. W. R. Martin
" P. Franklin
" N. Wallman
" G. B. Thomas
" R. Taylor
" Y. W. Lee

WSL:hms

MEMORANDUM NO. 94

Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

TO: J. W. Forrester, Robert R. Everett, Professor Franklin — 6346

FROM: W. S. Loud —

SUBJECT: Suggestions for Further Work

DATE: August 6, 1947

There are several directions in which I feel that mathematical research in connection with the project might well be carried forward. I am sure that Professor Franklin, with his wider experience will be able to select those which are of value. I might make the suggestion that people of Mr. Reich's caliber be given their own head for a while. I think they will find important subjects for investigation given the freedom of investigation.

I have some remarks to make about specific topics. They are of necessity incomplete. The work already done is in my notebooks and Miss Florencourt will know about it.

I. The constant energy method for solving differential equations.

We did quite a bit of work on this investigation. Precise quantitative results were not obtained owing to the mathematical complexity involved. I think the idea will be of value in physical problems because it takes into account such an important physical fact. It is particularly of value in non-dissipative systems.

The principal advantage of the method seems to be that it does not introduce bad amplitude distortion. The method of polynomial extrapolation and integration inevitably introduces this. The frequency and phase distortions are no worse than the other method. Also it is a relatively simple procedure from the point of view/set-up, at least for a small number of equations. It may become too cumbersome for the entire ASCA problem.

Under disadvantages I could list first of all the fact that square roots have to be taken and some criterion for the proper algebraic sign selected. This

might involve considerable difficulty. Mr. Reich has some ideas which may help on this score. The method of energies although reasonably simple may turn out to be too complex for complicated problems. The question of damped vibrations is one that has not been touched on. I have no knowledge of how the method could be applied there.

II Algebraic Equations

Some work was done on the subject of algebraic equations. For the case of simultaneous linear equations I expect that the work done at Princeton is better than could be done here. I think that attack on other problems, using their results as much as possible would be the best plan.

The next most complicated problem is that of simultaneous quadratic equations. The literature does contain much information on number of real roots etc., but complete analysis of susceptibility to numerical solution does not seem to have been carried through. We have done some work on iterative procedures, and the results are mathematically interesting though they do not promise much for machine computation as yet. I believe that statisticians are investigating this subject and may have some useful results. Methods other than iterative might also be studied in these cases.

III Error Analysis

The knowledge of error occurring in the numerical solution of a differential equation is very important as we well know. Acting on Dr. Goldstine's suggestion we carried through an analysis of error based on known results from the theory of differential equations. The results are only of academic interest. The generality of application of the theorems makes them very crude of necessity. I think investigation here should go forward in two directions. First the analysis for error should be specialized. Sharper results with more restrictions on the differential equations to which they apply should be obtained. This idea is contrary to the current mathematical fashion of generalizations, but I believe it should be done. Second the subject of round-off error should be treated. I agree with Wiener that this must be handled statistically. I am not too sure that his predictions theory can be used without some essential modification, but I don't know. For more remarks on this see the accompanying report on my investigation of Wiener's ideas.

In general I think it would be a good idea if the philosophy of high-speed computers were kept in mind during mathematical investigations. This will have the effect of keeping to fairly concrete cases so that the results will not be overly abstract, and should also provide the proper motivations for work.

In conclusion I can say that my future work will deal for a time with these same subjects. I can let you know if I arrive at results which will be useful.

Warren S. Loud
W. S. Loud

WSL/rp

MEMORANDUM NO. M-95

Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

TO: J. W. Forrester, R. R. Everett, Professor Franklin — 6345

FROM: W. S. Loud — Page 1 of 4 pages

SUBJECT: Report on Wiener's Ideas

DATE: August 11, 1947

For the past three months my efforts have been turned quite completely to obtaining something concrete out of Prof. Wiener's ideas. The following is a report on what I have learned from this study with what conclusions I have been able to draw.

I. The Mathematical Ideas and Their Scope of Application

Professor Wiener has developed a statistical prediction theory built on Fourier Transforms. He has applied this theory to problems of filtering, differentiating, etc. His most complete results are in a linear theory. He has given an indication of a non-linear theory, but I cannot report on that because I do not understand it as yet.

The essential mathematical quantities and processes used are the auto-correlation and cross-correlation of functions and the operation of Fourier Transform. For example, let f(t) be a function of time. Let us consider the case of discrete times since this is most applicable for digital computation. We wish to find a formula for f(t+h) in terms of f(t), f(t-h), f(t-2h),... We assume a linear formula

$$f(t+h) = \sum_{n=0}^{\infty} a_n f(t-nh),$$

and we seek the coefficients a_n so that the mean-square error is a minimum. It turns out that the best coefficients are the Fourier coefficients of a function

$$k(\omega) = \sum_{n=0}^{\infty} a_n e^{i n \omega}$$

We can obtain $k(\omega)$ in terms of the auto-correlation of f(t).

Definitions:

$$\phi_n = \lim_{N\to\infty} \frac{1}{2N} \sum_{K=-N}^{N} f(t+Kh+nh)\ \overline{f(t+Kh)}$$

$$\Phi(\omega) = \sum_{n=-\infty}^{\infty} \phi_n e^{in\omega}$$

$\Psi(\omega)$ is a function such that

(a) $\left|\Psi(\omega)\right|^2 = \Phi(\omega)$

(b) $\Psi(\omega)$ has neither poles nor zeros for $\mathrm{Im}\ \omega \geq 0$.

Then,

$$k(\omega) = \frac{1}{2\pi\,\Psi(\omega)} \sum_{n=0}^{\infty} e^{in\omega} \int_{-\pi}^{\pi} \Psi(u) e^{-i(n+1)u}\ du$$

We find a_n by

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} k(\omega) e^{-in\omega}\ d\omega$$

If we wished, $f(t+\lambda h)$ where λ is an integer, the corresponding $k(\omega)$ would be

$$\frac{1}{2\pi\,\Psi(\omega)} \sum_{n=0}^{\infty} e^{in\omega} \int_{-\pi}^{\pi} \Psi(u) e^{-i(n+\lambda)u}\ du$$

Other problems that can be solved are differentiation and filtering. The case of differentiation is covered in Memorandum M-78 for the continuous case. For the discrete case where we would wish

$$f'(t) = \sum_{n=0}^{\infty} a_n f(t - nh),$$

$$k(\omega) = \sum_{n=0}^{\infty} a_n e^{in\omega}$$

would be given by

$$k(\omega) = \frac{1}{2\pi \overline{Y}(\omega)} \sum_{n=0}^{\infty} e^{in\omega} \int_{-\pi}^{\pi} (-iu)\, Y(u) e^{-inu}\, du$$

For the case of filtering we have $f(t) = f_1(t) + f_2(t)$ where $f(t)$ is the input and $f_1(t)$ is the desired message and $f_2(t)$ is noise. We wish

$$f_1(t+\lambda h) = \sum_{n=0}^{\infty} a_n f(t-nh)$$

that is the message output in terms of the past of the input. Again, if

$$k(\omega) = \sum_{n=0}^{\infty} a_n e^{in\omega},$$

we have

$$k(\omega) = \frac{1}{2\pi \overline{Y}(\omega)} \sum_{n=0}^{\infty} e^{in\omega} \int_{-\pi}^{\pi} \frac{\Phi_1(u)}{\overline{Y}(u)} e^{-i(n+\lambda)u}\, du$$

where we have

$$\Phi_n^{(1)} = \lim_{N\to\infty} \sum_{K=-N}^{N} f_1(t+nh+Kh)\ \overline{f_1(t+Kh)}$$

and

$$\Phi_1(\omega) = \sum_{-\infty}^{\infty} \Phi_n^{(1)} e^{in\omega}$$

This is making the not unreasonable assumption that message and noise are independent statistically.

Explanatary Remarks:

In the filtering formula $\overline{Y(u)}$ is the complex conjugate of $Y(u)$. As an example of this factoring of Φ, if

$$\Phi(\omega) = \frac{1+\omega^2}{4+\omega^2}$$

$$Y(\omega) = \frac{\omega+i}{\omega+2i}\ ;\quad \overline{Y(\omega)} = \frac{\omega-i}{\omega-2i}$$

Note that Y has no singularities or zeros for $\mathrm{Im}\ \omega \geq 0$.

It is assumed in all of these formulas that there is enough statistical information about the function f(t) so that $\Phi(\omega)$, which depends on the future of f as well as its past, is known.

II. Limitations on Application

The application of the above theory is limited for reasons which are stated by Professor Wiener in his paper on Prediction Theory. The difficulty begins to show itself when we search for a concrete example.

The function $\Phi(\omega)$ turns out to be identically zero or non-existent in most cases that occur to mind readily. It turns out in fact that if $\Phi(\omega)$ is to be finite and not identically zero, the frequency spectrum of f(t) must be everywhere continuous, but nowhere differentiable. This implies that the variation of f must be extremely random. Wiener says definitely that his method is not of use with analytic functions, for these are functions whose values for all time are determined by small intervals. This fact would seem seriously to limit the method for solving the usual type of differential equation.

In fact I would say that for improving truncation error alone, the method is not of use. For round-off, which has this random character, the filtering idea may well work if the time solution is also somewhat random.

III. Conclusions

1) Much more work must be done before the ideas presented will find application for digital computation. Their best place is in minimization of round-off error.

2) It may well be that the non-linear theory will find its use more easily.

3) Concrete examples will be hard to construct. If f(t) is reasonably simple, $\Phi(\omega)$ is useless. If $\Phi(\omega)$ is at all simple, f(t) is a hopelessly messy function.

WARREN S. LOUD

WSL:has
c. M.I.Florencourt
E. Reich

Memorandum M-100

Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

To: Jay W. Forrester, R. R. Everett,
M. I. Florencourt, Professor Franklin

6345

From: Edgar Reich

Subject: M.A.A. Symposium on Computing Machines, Yale University,
September 2, 1947, 10 A.M. - 12 N.

Date: September 3, 1947

The symposium consisted of two addresses, the first by H. H. Aiken of Harvard, the second by John von Neumann of the Institute for Advanced Study, followed by a discussion.

Aiken described the mechanical components of the Harvard machines. Most of this material including the slides shown is already available in the "Manual of Operation". Von Neumann's talk contained remarks which might be of interest to the Laboratory, and excerpts from it, as well as from the discussion that followed, are summarized below.

Up to the present no significant success has been achieved in non-linear theory. The types of singularities encountered in non-linear systems are not like those of linear systems, and at the present stage of the art their character can, in most cases, only be obtained by physical experiment and/or extensive computing. The data obtained by the machine will enable us to examine these systems, add will thus provide an intuitive guide for a rigorous mathematical attack.

Dr. von Neumann compared analog to digital machines, in so far as accuracy is concerned. When two numbers are multiplied by an analog method, there is an error introduced due to the noise level of the analogous system. This noise level is of an essentially statistical nature. Physical limitations prevent an indefinite increase of the accuracy. We cannot hope for an accuracy better than 1 part in 100,000 with the analog type of computer. In digital machines, we also have a noise, caused, in this case, by the round-off procedure. With this type of computer, however, the effect of the noise can be made arbitrarily small by simply increasing the number of digits the machine handles (i.e. increasing the register length).

The development of efficient techniques for coding problems for the digital computer will make an extension of the field covered by mathematical logic necessary. Today mathematical logic requires nothing more of a proof than that it be valid, that is, consist of at most a finite number of valid deductions. No attention has been paid to the actual number of deductions, as long as this is finite. A proof, or construction, containing say five steps is for logical purposes considered no better than one made up of five hundred. In the coding of problems for digital computers, on the other hand, it is desirable to either minimize the total number of steps, or to minimize the total number of times a specified operation (such as, for instance, division) is used in a particular construction. This problem will have to be investigated in order to make it possible to decide what operations are to be built into a machine, and in order to use the storage capacity most efficiently.

In answer to a question as to whether or not the computer could be used to attack problems of number theory by brute force (for instance, whether it could be used to sift through enormous sets of numbers in search for a number with some special property), Dr. von Neumann replied that this application would be unlikely. The machine provides a time-reduction factor of about 10^6 over human computation, but the number of elements encountered in the aforementioned sets are often of astronomical magnitude, in which case the time required might still be too long.

During the discussion, someone mentioned the fact that the computer should be able to solve certain "highly important" non-linear difference equations arising in the Theory of Evolution.

Edgar Reich

Edgar Reich

ER:vh

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: LOCATION OF TARGET FROM COMBINED OBSERVATIONS

TO: Jay W. Forrester, R. R. Everett

FROM: Philip Franklin

DATE: October 21, 1947

The range and direction of a target T is observed on several stations whose location is known. From these observations we wish to determine the most probable position of the target. Assuming that the direction information has a precision at least comparable with the range information, the following procedure leads to a determination by an application of the theory of least squares.

As a first case, restrict matters to a plane in which we have two stations, S_1 and S_2. Introduce some fixed (xy) coordinate system. Let $S_1 = (x_1, y_1)$. Put $R_1 = S_1T$, the observed range from S_1. Put A_1 = azimuth of S_1T measured from the direction of OX, the observed bearing from S_1. Then the value of A_1 restricts T to a straight line whose equation is

$$(x - x_1) \sin A_1 - (y - y_1) \cos A_1 = 0, \text{ or } L_1 = 0, \qquad (1)$$

where L_1 equals the linear expression in the left member. If the probable error of A_1 is dA_1, the probable error of distance from line (1) near T is $R_1 dA_1$. Assign the squared reciprocal of this, $w_1 = 1/(R_1 dA_1)^2$, as a weight or measure of precision for the Eq. (1).

The value of R_1 restricts T to a circle which we approximate by its tangent at T. This straight line has as its equation

$$(x - x_1) \cos A_1 + (y - y_1) \sin A_1 - R_1 = 0, \text{ or } L_2 = 0. \qquad (2)$$

The probable error of distance from line (2) near T is dR_1, the probable error of R_1. Assign the squared reciprocal of this, $w_2 = 1/(dR_1)^2$ as a weight or measure of precision for the Eq. (2).

Similarly the observations R_2, A_2 for station S_2 and their probable errors dR_2, dA_2 lead to two equations $L_3 = 0$, with a weight w_3 and $L_4 = 0$ with a weight w_4. The expressions for L_3 and w_3 are like those for L_1 and w_1, except that subscripts 1 are changed to 2. Similarly, the expressions for L_4 and w_4 are like those for L_2 and w_2.

According to the theory of normal errors, the probability of T being at a given position (x,y) is equal to Ke^{-kE} where

$$E = w_1 L_1^2 + w_2 L_2^2 + w_3 L_3^2 + w_4 L_4^2.$$

The probability is greatest when this is least, and the normal equations for x and y are

$$\begin{aligned} w_1 \sin A_1 L_1 + w_2 \cos A_1 L_2 + w_3 \sin A_2 L_3 + w_4 \cos A_2 L_4 &= 0 \\ -w_1 \cos A_1 L_1 + w_2 \sin A_1 L_2 - w_3 \cos A_2 L_3 + w_4 \sin A_2 L_4 &= 0. \end{aligned} \tag{4}$$

The solution of these two linear equations x_T, y_T is to be taken as the target position. It is the center of a family of similar ellipses, E = constant, where E is the expression in (3). All points on any one ellipse are equally likely positions for T.

EXAMPLE 1: Let $w_2 = w_1$ and $w_4 = w_3$. Here the observations at S_1 alone lead to a circular probability distribution with center at a point we call T_1, and similarly those at S_2 alone lead to circles with center at a point we call T_2. The ellipses E in this case reduce to circles, and the point x_T, y_T found from Eq. (4) is on the line T_1T_2 with $T_1T/TT_2 = w_3/w_1$.

EXAMPLE 2: Let w_1 and w_3 be very large compared with either w_2 or w_4. Here x_T, y_T will be very near the intersection of L_1 and L_3, the position determined by the bearings alone, and slightly biassed toward the lines L_2 and L_4.

These two examples show that the method leads to plausible results in simple cases.

The General Problem in a Plane: For n stations S_i, $i = 1, 2, 3, \ldots n$, there are 2n linear expressions L_j, $j = 1, 2, 3, \ldots 2n$ and equations $L_j = 0$ with weights w_j. For j odd, $j = 2i - 1$, these are formed from the azimuths A_i similar to Eq. (1). For j even, $j = 2i$, these are formed from the ranges R_i similar to Eq. (2). Here the quantity E, such that E=constants gives ellipses of equal probability, is

$$E = \sum_{j=1}^{2n} w_j L_j^2 . \qquad (5)$$

$$\sum_{i=1}^{n} (w_{2i-1} \sin A_i L_{2i-1} + w_{2i} \cos A_i L_{2i}) = 0$$
$$\sum_{i=1}^{n} (-w_{2i-1} \cos A_i L_{2i-1} + w_{2i} \sin A_i L_{2i}) = 0. \qquad (6)$$

As in Example 1, if each $w_{2i} = w_{2i-1}$, the observations for each S_i alone lead to a circular distribution, and x_T, y_T is the center of gravity of n weights w_{2i}, where w_{2i} is placed at the point determined by the observations at S_i alone, i.e., the intersection of the lines $L_{2i-1} = 0$ and $L_{2i} = 0$. The ellipses E = constant reduce to circles.

Three-Dimensional Case: Consider a station S_1 in space whose coordinates are x_1, y_1, z_1 with respect to some fixed xyz system. Let a target be located by its observed range $R_1 = S_1T$, azimuth of the vertical plane through S_1T measured from the xz plane, taken as A_1, and polar declination or colatitude B_1 equal to the angle from the direction of OZ to S_1T. Then the value of A_1 restricts T to a plane whose equation is $L_1 = 0$, where

$$L_1 = (x - x_1) \sin A_1 - (y - y_1) \cos A_1 . \qquad (7)$$

The weight w_1 is the square reciprocal of the probable error of distance from the plane, $w_1 = 1/(R_1 \sin B_1 dA_1)^2$.

The value of B_1 restricts T to the plane $L_2 = 0$, where

$$L_2 = (x - x_1) \cos A_1 \cos B_1 + (y - y_1) \sin A_1 \cos B_1 - (z - z_1) \sin B_1 . \qquad (8)$$

The probable error of distance is $R_1\, dB_1$, so that the weight $w_2 = 1/(R_1\, dB_1)^2$.

The value of R restricts T to the plane $L_3 = 0$, where

$$L_3 = (x - x_1)\cos A_1 \sin B_1 + (y - y_1)\sin A_1 \sin B_1 + (z - z_1)\cos B_1 - R_1 \qquad (9)$$

The probable error of distance is dR_1, and the weight $w_3 = 1/(dR_1)^2$.

If there are n stations S_i, for each station, we can form three expressions L_{3i-2}, L_{3i-1}, and L_{3i} like L_1, L_2, L_3 and corresponding weights like w_1, w_2, w_3. With revised notation, we have 3n linear expressions L_j and equations $L_j = 0$ with weights w_j, where

$$L_j = a_j x + b_j y + c_j z - d_j. \qquad (10)$$

The probability that T is at x,y,z is now $P = K e^{-kE}$, where

$$E = \sum_{j=1}^{3n} w_j L_j^2 . \qquad (11)$$

E = constant gives the ellipsoids of equal probability. And the value x_T, y_T, z_T which makes P a maximum, makes E a minimum and satisfies the normal equations

$$\sum_{j=1}^{3n} w_j a_j L_j = 0$$

$$\sum_{j=1}^{3n} w_j b_j L_j = 0 \qquad (12)$$

$$\sum_{j=1}^{3n} w_j c_j L_j = 0.$$

Examples: If each $w_{3i} = w_{3i-1} = w_{3i-2}$, the observations for each S_i alone lead to a spherical distribution. The ellipsoids E = constant reduce to spheres. The point x_T, y_T, z_T, is the center of gravity of n weights w_{3i} where w_{3i} is placed at the point determined by the observations at S_i alone, i.e., the intersection of the three planes

$$L_{3i-2} = 0,\ L_{3i-1} = 0,\ L_{3i} = 0.$$

As a second example, let there be just two stations, and let

w_3 and w_6 be much smaller than the other weights. Also let $w_2 = w_1$ and $w_5 = w_4$. Then x_T, y_T, z_T as found from the normal equations (12) will be very near the common perpendicular to the lines S_1P_1 and S_2P_2 obtained by using directions only. And if the common perpendicular is P_1P_2, it will be near T, where $P_1T/TP_2 = w_4/w_1$.

The method again leads to reasonable results in these simple cases.

Calculation in the Plane: In the plane case, it might simplify the computation to replace the coefficients $\sin A_1$, $\cos A_1$ by $\tan A_1$, 1 or 1, $\cot A_1$. This is theoretically not correct, but since the weights are known only approximately, if we always changed the biggest of sine or cosine to unity, we would be only changing the weights by a factor between 0.5 and 1, which might not be too serious.

Philip Franklin

Philip Franklin

PF:has